

MENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A229 502

estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20543.

REPORT DATE

November 8, 1990

3. REPORT TYPE AND DATES COVERED

Final Report, 15 Sep 89 to 14 Sep 90

4. TITLE AND SUBTITLE

MULTIVARIATE MULTISCALE ANALYSIS

5. FUNDING NUMBERS

AFOSR-89-0524

61102F 2304/ A9

6. AUTHOR(S)

W. R. Madych

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Connecticut
Department of Mathematics, U-9
Storrs, CT 06269

8. PERFORMING ORGANIZATION
REPORT NUMBER

AFOSR-TR-90-1159

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NM
Bldg 410
Bolling AFB DC 20332-6448

10. SPONSORING / MONITORING
AGENCY REPORT NUMBER

AFOSR-89-0524

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The PI produced five (5) published papers during the period of support. The work centered on questions of Fourier/wavelet analysis. This is an area which promises to have a major impact on signal processing and numerical analysis and requires careful theoretical underpinning. Dr. Madych has resolved several questions related to translation invariance, multiscale analysis and Radon transforms.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION
OF REPORT

UNCLASSIFIED

18. SECURITY CLASSIFICATION
OF THIS PAGE

UNCLASSIFIED

19. SECURITY CLASSIFICATION
OF ABSTRACT

UNCLASSIFIED

20. LIMITATION OF ABSTRACT

SAR

MULTIVARIATE MULTISCALE ANALYSIS

Final Technical Report for AFOSR-89-0524

November 8, 1990

PI - W. R. Madych

AFOSR-TR- 90 1159

The following is a list of work completed while the principal investigator was partially supported by this grant.

The following are either published or accepted for publication:

1. Translation invariant multiscale analysis, pages 455-462 in *Recent Advances in Fourier Analysis and Its Applications*, J. S. Byrnes and J. L. Byrnes, eds., NATO ASI Series C, Vol. 315, Kluwer, Dordrecht, 1990.
2. Polyharmonic Splines, Multiscale Analysis, and Entire Functions, pages 205- 216 in *Multivariate Approximation and Interpolation*, W. Haussmann and K. Jetter, eds., International Series of Numerical Mathematics, Vol. 94, Birkhauser Verlag, Basel, 1990.
3. Summability and Approximate Reconstruction from Radon Transform Data, to appear in *Contemp. Math.*, 34 pp.

The following are preprints submitted for publication and technical reports:

1. Bounds on Multivariate Polynomials and Exponential Error Estimates for Multiquadric Interpolation, BRC/Math-TR-90-1, 19pp.
2. Error Estimates for Interpolation by Generalized Splines, BRC/Math-TR-90-3, 8 pp.

A copy of each of the above works is included in the enclosed volume.

This grant provided at least part of the cost of participation in each of the following conferences.

1. June '90 - Chamonix, France. Invited participant in conference on multivariate approximation and interpolation. Gave talk on error estimates for generalized splines.
2. August '90 - Oberwolfach, W. Germany. Invited participant in conference on the Radon Transform and its applications. Gave talk on wavelets and their potential applications in this area.

#157
Contents:

- 1) Translation invariant multiscale analysis; pages 455-462 in *Recent Advances in Fourier Analysis and Its Applications*, J. S. Byrnes and J. L. Byrnes, eds., NATO ASI Series C, Vol. 315, Kluwer, Dordrecht, 1990.
- 2) Polyharmonic Splines, Multiscale Analysis, and Entire Functions, pages 205- 216 in *Multivariate Approximation and Interpolation*, W. Haussmann and K. Jetter, eds., International Series of Numerical Mathematics, Vol. 94, Birkhauser Verlag, Basel, 1990.
- 3) Summability and Approximate Reconstruction from Radon Transform Data, to appear in *Contemp. Math.*, 34 pp
- 4) Bounds on Multivariate Polynomials and Exponential Error Estimates for Multiquadric Interpolation, BRC/Math-TR-90-1, 19pp.
- 5) Error Estimates for Interpolation by Generalized Splines, BRC/Math-TR-90-3, 8 pp.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



L (KR)
↑

Translation Invariant Multiscale Analysis

W. R. Madych*

Abstract

The notion of multiscale analysis introduced by R. R. Coifman and Y. Meyer is considered and the translation invariant case is characterized.

1 Introduction

Recall that a *dyadic multiscale analysis* of $L^2(R^n)$ is an increasing sequence $\mathcal{V} = \{V_j : j = \dots, -1, 0, 1, 2, \dots\}$ of closed subspaces of $L^2(R^n)$ which has the following properties:

- (1). $\bigcup_{j=-\infty}^{\infty} V_j$ is dense in $L^2(R^n)$ and $\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$.
- (2). $f(x)$ is in V_j if and only if $f(2x)$ is in V_{j+1} .
- (3). There is a lattice Γ in R^n such that for every f in V_0 and every γ in Γ the function f_γ is in V_0 . Here and in what follows we use the notation $f_\gamma(x) = f(x - \gamma)$.
- (4). There are two positive constants $C_2 \geq C_1 > 0$ and a function g in V_0 such that V_0 is the closed linear span of g_γ , $\gamma \in \Gamma$, and

$$C_1^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2 \leq \int_{R^n} \sum_{\gamma \in \Gamma} |a_\gamma g_\gamma(x)|^2 dx \leq C_2^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2.$$

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Preliminary report of work partially supported by a grant from the Air Force Office of Scientific Research.

An introduction to the subject may be found in [1,2]. A basic property of a multiscale analysis \mathcal{V} is the following:

- (5). There is a function ϕ in V_0 such that the collection $\{\phi_\gamma\}_{\gamma \in \Gamma}$ is an orthonormal basis in V_0 .

This fact may be regarded as a substitute for (4) and plays an important role in what follows.

A dyadic multiscale analysis is *translation invariant* if all the translates of f , $\{f_y : y \in R^n\}$, are in V_0 whenever f is in V_0 .

The canonical example of a translation invariant multiscale analysis of $L^2(R)$ is when V_0 is the collection of those functions in $L^2(R)$ whose Fourier transforms are supported in the interval $[-\pi, \pi]$. A natural choice of ϕ in this case is given by

$$\phi(x) = \frac{\sin \pi x}{\pi x}.$$

The point of this paper is to give a characterization of translation invariant multiscale analyses. For the sake of clarity in what follows we will restrict our attention to the case $n = 1$ and $\Gamma = Z$, the lattice of integers. The statements and arguments in the general case are completely analogous to this basic case.

We now briefly digress to list some of the conventions which are used here: The Fourier transform \hat{f} of a function f is defined by

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} f(x) dx$$

whenever it makes sense and distributionally otherwise. Basic facts concerning Fourier transforms and distributions will be used without further elaboration in what follows. To avoid the pedantic repetition of "almost everywhere" and other modifying phrases which are inevitably necessary when dealing with functions defined almost everywhere, all equalities between functions and other related notions are interpreted in the distributional sense whenever possible. The term support is also used in the distributional sense; in particular the support of a function f in $L^2(R)$ is a well defined closed set. If W is a collection of tempered distributions then \widehat{W} is the collection of Fourier transforms of elements of W , in other words $\widehat{W} = \{f : f = \hat{g} \text{ for some } g \text{ in } W\}$. For a subset Ω of R and a real number

r the sets $r\Omega$ and $\Omega + r$ are defined by $r\Omega = \{x : x = r\omega \text{ for some } \omega \text{ in } \Omega\}$ and $\Omega + r = \{x : x = \omega + r \text{ for some } \omega \text{ in } \Omega\}$; $L^2(\Omega)$ is the L^2 closure of the subspace of those functions in $L^2(R)$ whose support is contained in Ω . For notational simplicity we use Q to denote the closed interval $[-\pi, \pi]$.

We can now conveniently state our main observation.

Theorem *Suppose \mathcal{V} is a translation invariant dyadic multiscale analysis of $L^2(R)$. Then $\hat{V}_0 = L^2(\Omega)$ where Ω is a closed subset of R which has the following properties:*

- (a). $\Omega \subset 2\Omega$.
- (b). $\{\Omega + 2\pi j\} \cap \{\Omega + 2\pi k\}$ is a set of Lebesgue measure 0 for any pair of integers such that $j \neq k$.
- (c). $\bigcup_{k=-\infty}^{\infty} \{\Omega + 2\pi k\} = R$.
- (d). $\bigcup_{k=1}^{\infty} L^2(2^k\Omega)$ is dense in $L^2(R)$.

Conversely, if V_k , $k \in Z$ is defined by $\hat{V}_k = L^2(2^k\Omega)$ where Ω is a closed subset of R which satisfies the properties above then the sequence of subspaces $\{V_k\}$ is a translation invariant multiscale analysis of $L^2(R)$.

Remark 1 In view of the example given above it is very tempting to conjecture that the set Ω in the Theorem must be of the form $\Omega = Q + \alpha$ for some real number α which satisfies $\pi < \alpha < \pi$. Certainly such Ω 's satisfy the desired conditions. However the conditions of the Theorem are satisfied by Ω 's which need not be connected as the following example due to Rudi Lorentz shows.

$$\Omega = \left[-\frac{5\pi}{4}, -\pi\right] \cup \left[-\frac{3\pi}{4}, \frac{3\pi}{4}\right] \cup \left[\pi, \frac{5\pi}{4}\right] .$$

Remark 2 Consider

$$\begin{aligned}\Omega_a &= [-1, 1] \cup [2\pi + 1, 4\pi - 1] \\ \Omega_b &= [-5, 5] \\ \Omega_c &= [-1, 1] \\ \Omega_d &= [0, 2\pi] .\end{aligned}$$

It is not difficult to verify that each Ω_α listed above is a closed set which fails to satisfy condition (a) but satisfies the remaining conditions in the Theorem. These examples show that conditions (a)-(d) are not redundant.

Remark 3 Note that condition (a) implies that 0 is contained in Ω . In addition to this it is clear that if Ω contains a neighborhood of the origin then it satisfies condition (d). In view of this it seems reasonable to suspect that subsets Ω which satisfy the conditions of the Theorem must contain an open neighborhood of the origin. That this is not the case can be seen by considering the following example of Ω :

$$\left\{ \bigcup_{k=1}^{\infty} [-(2 - 2^{-k})2^{-k}\pi, -2^{-k}\pi] \right\} \cup [0, \pi] \left\{ \bigcup_{k=1}^{\infty} [(2 - (2 - 2^{-k})2^{-k})\pi, (2 - 2^{-k})\pi] \right\}.$$

Remark 4 In view of the examples listed above it may be of some interest to obtain a significantly more lucid description of the set Ω than that given in the Theorem.

A corollary concerning wavelets generated by \mathcal{V} is recorded at the end of Section 2.

2 Details

We begin by establishing a basic lemma. First recall that the indicator function of a set Ω is usually denoted by χ_Ω and satisfies

$$\chi_\Omega(\xi) = \begin{cases} 1 & \text{if } \xi \in \Omega \\ 0 & \text{otherwise.} \end{cases}$$

Lemma *Suppose \mathcal{V} is a translation invariant dyadic multiscale analysis of $L^2(R)$ and ϕ is a function whose existence is guaranteed by (5). Then*

$$|\hat{\phi}| = \chi_\Omega$$

where χ_Ω is the indicator function of a closed set Ω which has properties (a)-(d) in the statement of the Theorem.

Let Ω be the support of ϕ . To prove the lemma we will first show that Ω satisfies property (b).

Recall that (5) implies that for all f in V_0 we may write

$$(6) \quad \hat{f}(\xi) = g(\xi)\hat{\phi}(\xi)$$

where g is 2π periodic and square integrable over Q . In particular, since V_0 is translation invariant, ϕ_y is in V_0 so setting $\alpha = -y$ we may write

$$e^{i\alpha\xi}\hat{\phi}(\xi) = g(\xi)\hat{\phi}(\xi)$$

for some such g . Hence

$$e^{i\alpha(\xi-2\pi m)}\hat{\phi}(\xi-2\pi m) = g(\xi-2\pi m)\hat{\phi}(\xi-2\pi m) = g(\xi)\hat{\phi}(\xi-2\pi m)$$

which implies that

$$e^{i\alpha(\xi-2\pi m)} = g(\xi)$$

on $\Omega + 2\pi m$. For two different values of m the last equality implies that

$$e^{i\alpha(\xi-2\pi j)} = e^{i\alpha(\xi-2\pi k)}$$

on $\{\Omega + 2\pi j\} \cap \{\Omega + 2\pi k\}$. Re-expressing the last relation as

$$e^{i\alpha(\xi-2\pi k)}(e^{i2\pi\alpha(k-j)} - 1) = 0$$

it is clear that either α is an integer, j is equal to k , or $\{\Omega + 2\pi j\} \cap \{\Omega + 2\pi k\}$ is a set of measure zero. Since α may be any real number we conclude that $\{\Omega + 2\pi j\} \cap \{\Omega + 2\pi k\}$ has measure zero whenever $j \neq k$.

Now, since $\hat{\phi}_k(\xi) = e^{-ik\xi}\hat{\phi}(\xi)$, $k \in Z$, are orthonormal, we may write

$$(7) \quad \int_R \phi_k(x)\overline{\phi_\ell(x)}dx = \int_R e^{im\xi}|\hat{\phi}(\xi)|^2d\xi = \int_Q e^{im\xi} \sum_{j \in Z} |\hat{\phi}(\xi-2\pi j)|^2dx = \begin{cases} 1 & \text{when } m = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $m = \ell - k$. The last equality implies that

$$\sum_{j \in Z} |\hat{\phi}(\xi-2\pi j)|^2 = \frac{1}{2\pi}$$

on R and since $\{\Omega + 2\pi j\} \cap \{\Omega + 2\pi k\}$ has measure zero whenever $j \neq k$ we may conclude that

$$|\hat{\phi}(\xi)| = \frac{1}{\sqrt{2\pi}} \chi_{\Omega}(\xi)$$

and

$$\bigcup_{k \in \mathbb{Z}} \{\Omega + 2\pi k\} = R.$$

To see (a) observe that (2) and the facts demonstrated above imply that

$$\chi_{\Omega}(\xi) = h(\xi) \chi_{\Omega}(\xi/2)$$

where h is 4π periodic and square integrable over $2Q$. Since $\chi_{2\Omega}(\xi) = \chi_{\Omega}(\xi/2)$, the last equality involving h implies that χ_{Ω} vanishes whenever $\chi_{2\Omega}$ does so $\Omega \subset 2\Omega$.

Finally, the fact that $\bigcup_{k=1}^{\infty} L^2(2^k \Omega)$ is dense in $L^2(R)$ is an immediate consequence of property (1). The proof of the Lemma is complete.

Now, suppose ϕ and Ω are as in the Lemma and its proof. Since V_0 consists of functions f which satisfy (6) it is clear that \hat{V}_0 is contained in $L^2(\Omega)$.

To see that $\hat{V}_0 = L^2(\Omega)$ let f be any element in $L^2(\Omega)$ and let h be defined by

$$h(\xi) = \begin{cases} \overline{\hat{\phi}(\xi)} / |\hat{\phi}(\xi)| & \text{if } \hat{\phi}(\xi) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

By virtue of the properties of Ω established above it is clear that

$$\frac{1}{\sqrt{2\pi}} \chi_{\Omega}(\xi) = \left(\sum_{j \in \mathbb{Z}} h(\xi - 2\pi j) \right) \hat{\phi}(\xi)$$

and

$$\hat{f}(\xi) = g(\xi) \hat{\phi}(\xi)$$

where

$$g(\xi) = \sqrt{2\pi} \left(\sum_{j \in \mathbb{Z}} \hat{f}(\xi - 2\pi j) \right) \left(\sum_{j \in \mathbb{Z}} h(\xi - 2\pi j) \right)^{-1}.$$

Thus f satisfies (6) and hence we may conclude that $L^2(\Omega)$ is contained in \hat{V}_0 .

The Lemma together with the last observation imply the first assertion of the Theorem.

To see the converse, let \mathcal{V} be the sequence of subspaces $\{V_k\}$, $k \in Z$, defined by $\widehat{V}_k = L^2(2^k\Omega)$ where Ω is a closed set which satisfies properties (a)-(d) of the Theorem.

The fact that \mathcal{V} is translation invariant and, in particular, satisfies property (3) with $\Gamma = Z$ is an immediate consequence of the definition. Property (2) is also immediate. That \mathcal{V} is an increasing sequence of subspaces and $\bigcup_{k \in Z} V_k$ is dense in $L^2(R)$ are consequences of properties (a) and (d).

That $\bigcap_{k \in Z} V_k = \{0\}$ follows from the fact that the measure of Ω is finite. Indeed, its measure is 2π which can be seen from

$$\int_R \chi_\Omega(\xi) d\xi = \sum_{j \in Z} \int_{Q+2\pi j} \chi_\Omega(\xi) d\xi = \int_Q \sum_{j \in Z} \chi_\Omega(\xi - 2\pi j) d\xi = 2\pi$$

by using properties (b) and (c).

Finally, to see property (5) take

$$\hat{\phi} = \frac{1}{\sqrt{2\pi}} \chi_\Omega$$

and use properties (b) and (c) to write (7) which shows that $\phi_k(x)$, $k \in Z$, are orthonormal and, for f in V_0 ,

$$\hat{f}(\xi) = \sqrt{2\pi} \left(\sum_{k \in Z} \hat{f}(\xi - 2\pi k) \right) \hat{\phi}(\xi)$$

or

$$f(x) = \sum_{k \in Z} f(j) \phi(x - k)$$

which shows that they are complete in V_0 .

This completes the proof of the Theorem.

Remark 5 Suppose \mathcal{V} is a translation invariant multiscale analysis and Ω is a closed set such that $\widehat{V}_0 = L^2(\Omega)$. Then if W_0 is the orthogonal complement of V_0 in V_1 , $\widehat{W}_0 = L^2(\Upsilon)$ where $\Upsilon = 2\Omega \setminus \Omega$. Let ψ be such that the set $\{\psi_k : k \in Z\}$ is an orthonormal basis for W_0 . Such a ψ may be referred to

as a wavelet. Using reasoning analogous to the proof of the Theorem, it is clear that ψ is a wavelet if and only if

$$|\hat{\psi}| = \frac{1}{\sqrt{2\pi}} \chi_{\mathbb{R}}.$$

Now, an analyzing wavelet in the sense of Meyer, [1], is globally integrable and hence its Fourier Transform must be continuous. Clearly ψ is not such an analyzing wavelet.

Corollary *A translation invariant multiscale analysis cannot give rise to analyzing wavelets in the sense of Meyer.*

References

- [1] J. M. Combs, A. Grossman, and Ph. Tchamitchian, eds., *Wavelets*, Proceedings of the International Conference in Marseille, 1987, Springer-Verlag, Berlin, 1989.
- [2] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, 41 (1988), 909-996.

Polyharmonic Splines, Multiscale Analysis, and Entire Functions

W. R. Madych*

Abstract

Two applications of cardinal polyharmonic splines are considered. First, the notion of multiscale analysis in the sense of R. Coifman and Y. Meyer is defined and it is shown that certain subclasses of polyharmonic splines generate such analyses. Next, it is shown that polyharmonic splines can be used in a summability method for the recovery of a large class of entire functions of exponential type from samples taken on the integer lattice.

1 Introduction

If k is a positive integer an n -variate k -harmonic cardinal spline is a tempered distribution f on R^n such that $\Delta^k f$ is a measure supported on the integer lattice Z^n in R^n . Symbolically

$$(1) \quad \Delta^k f(x) = \sum_{j \in Z^n} a_j \delta(x - j)$$

where Δ is the n -variate Laplacian, $\Delta^k = \Delta \Delta^{k-1}$ for $k > 1$, and $\delta(x)$ is the unit Dirac measure supported at the origin. The class of k -harmonic cardinal splines on R^n is denoted by $SH_k(R^n)$. A *polyharmonic spline* is one which is k -harmonic for some k .

The basic properties of these splines are recorded in [12, 13]. Here we are primarily concerned with certain applications. More specifically, in Section 2 we show that certain subclasses of the spaces $SH_k(R^n)$ provide examples of multiscale analyses whose general theory was developed recently by Y. Meyer and his collaborators. In Section 3 we show that polyharmonic splines provide a summability method for the recovery of entire functions of exponential type from samples on the lattice Z^n . The introductions to these sections contain a more complete description of their contents. Various comments which are

*Department of Mathematics, University of Connecticut, Storrs, CT 06268. Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-89-0524

not particularly germane to the development but are pertinent to the subject matter are given in Section 4.

The notation used here is standard, if necessary see [13] for a more detailed explanation. Here we merely remind the reader that there are several common normalizations for the Fourier transform. In this note we use

$$\hat{\psi}(\xi) = (2\pi)^{-n/2} \int_{R^n} \psi(x) e^{-i\langle \xi, x \rangle} dx$$

for the Fourier transform $\hat{\psi}$ of a test function ψ . Also, for convenience, we use the notation $T(\xi)f(\xi)$ for the product of a distribution T and a function f even when such an expression makes no sense pointwise; however, its meaning in the distributional sense should be clear from the context.

2 Multiscale analysis

Recall that a *dyadic multiscale analysis* of $L^2(R^n)$ is an increasing sequence $\mathcal{V} = \{V_j : j = \dots, -1, 0, 1, 2, \dots\}$ of closed subspaces of $L^2(R^n)$ which has the following properties:

- (i). $\bigcup_{j=-\infty}^{\infty} V_j$ is dense in $L^2(R^n)$ and $\bigcap_{j=-\infty}^{\infty} V_j = \{0\}$.
- (ii). $f(x)$ is in V_j if and only if $f(2x)$ is in V_{j+1} .
- (iii). There is a lattice Γ in R^n such that for every f in V_0 and every γ in Γ the function f_γ is in V_0 . Here and in what follows we use the notation $f_\gamma(x) = f(x - \gamma)$.
- (iv). There are two positive constants $C_2 \geq C_1 > 0$ and a function g in V_0 such that V_0 is the closed linear span of g_γ , $\gamma \in \Gamma$, and

$$C_1^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2 \leq \int_{R^n} \sum_{\gamma \in \Gamma} |a_\gamma g_\gamma(x)|^2 dx \leq C_2^2 \sum_{\gamma \in \Gamma} |a_\gamma|^2.$$

An introduction to the subject may be found in [5, 6, 9]. A basic property of a multiscale analysis \mathcal{V} is the following:

- (v). There is a function ϕ in V_0 such that the collection $\{\phi_\gamma\}_{\gamma \in \Gamma}$ is an orthonormal basis in V_0 .

This fact may be regarded as a substitute for (iv) and plays an important role in what follows.

Finally, we say that a dyadic multiscale analysis is *composed of generalized spline functions* in the sense of Y. Meyer if all the elements of the subspace V_0 are continuous and the mapping which maps f into the sequence of values $\{f(\gamma)\}_{\gamma \in \Gamma}$ is an isomorphism from V_0 onto $\ell^2(\Gamma)$.

In what follows we will restrict our attention to the case $\Gamma = \mathbb{Z}^n$, the lattice of integers. The statements and arguments in the general case are completely analogous to this basic case. Note that, by virtue of item (ii) in the above definition, a multiscale analysis is determined by the subspace V_0 .

For any positive integer k consider the subspace

$$V_0(k) = SH_k(\mathbb{R}^n) \cap L^2(\mathbb{R}^n).$$

For any integer m define

$$V_m(k) = \{f : f(2^{-m}x) \text{ is in } V_0(k)\}.$$

In other words, f is in $V_0(k)$ if and only if $f(2^m x)$ is in $V_m(k)$. The questions we address in this section are the following: For what values of k , if any, is the increasing family

$$(2) \quad \mathcal{V}(k) = \{V_m(k) \mid m = \dots, -1, 0, 1, 2, \dots\}$$

a dyadic multiscale analysis of $L^2(\mathbb{R}^n)$ and when is it composed of spline functions in the sense of Y. Meyer.

The answers to the above questions are not difficult. We begin by recalling several facts concerning $SH_k(\mathbb{R}^n)$.

If f is in $SH_k(\mathbb{R}^n)$ then for x in a sufficiently small neighborhood of \mathbf{j}

$$(3) \quad f(x) = a_{\mathbf{j}} E_k(x - \mathbf{j}) + h(x)$$

where E_k is the fundamental solution of Δ^k , namely,

$$(4) \quad E_k(x) = \begin{cases} c(n, k) |x|^{2k-n} & \text{if } n \text{ is odd} \\ c(n, k) |x|^{2k-n} \log |x| & \text{if } n \text{ is even,} \end{cases}$$

here the constant $c(n, k)$ depends only on n and k and is chosen so that $\Delta^k E_k(x) = \delta(x)$; h is infinitely differentiable and $a_{\mathbf{j}}$ is the constant in representation (1) of $\Delta^k f$. Moreover if $\Delta^k f = 0$ then f is a k -harmonic polynomial. These observations together with the behavior of E_k at the origin allow us to conclude the following:

Proposition 1 *If $4k \leq n$ then $V_0(k) = \{0\}$.*

If $4k > n$ then the function Φ_k defined by the formula for its Fourier transform

$$(5) \quad \hat{\Phi}_k(\xi) = (2\pi)^{-n/2} \frac{|\xi|^{-2k}}{\left(\sum_{\mathbf{j} \in \mathbb{Z}^n} |\xi - 2\pi \mathbf{j}|^{-4k}\right)^{1/2}}$$

is well defined and is in $V_0(k)$. Note that, by virtue of the same reasoning as used in [13], $\hat{\Phi}_k$ can be extended to a function which is holomorphic in

a tube containing R^n in complex n -space. As a consequence Φ_k enjoys the representation

$$\Phi_k(x) = \sum_{j \in Z^n} a_j E_k(x - j)$$

where the sequence $\{a_j\}$ decays exponentially, namely,

$$|a_j| \leq C e^{-\epsilon |j|}$$

and where C and ϵ are positive constants which depend only on k and n .

Observe that

$$\begin{aligned} \int_{R^n} \Phi_k(x - j) \overline{\Phi_k(x - m)} dx &= (2\pi)^{-n} \int_{R^n} \frac{e^{i(m-j, \xi)} |\xi|^{-4k}}{\sum_{j \in Z^n} |\xi - 2\pi j|^{-4k}} d\xi \\ &= (2\pi)^{-n} \int_{Q^n} e^{i(m-j, \xi)} d\xi \end{aligned}$$

where

$$Q^n = \{\xi = (\xi_1, \dots, \xi_n) : -\pi \leq \xi_j \leq \pi, j = 1, \dots, n\}$$

and hence the collection

$$(6) \quad \{\Phi_k(x - j) : j \in Z^n\}$$

is an orthonormal set in $V_0(k)$. Thus $V_0(k)$ certainly contains much more than the 0 element in this case. As it turns out, the collection (6) is also complete in $V_0(k)$.

To see this, let f be any element in $V_0(k)$. Viewing \hat{f} as a distribution on

$$S_0(R^n) = \{\phi \in S(R^n) : D^\nu \phi(0) = 0 \text{ for all multi-indices } \nu\},$$

it is clear that

$$(7) \quad \hat{f}(\xi) = P(\xi) |\xi|^{-2k}$$

where P is locally square integrable and periodic, that is, $P(\xi - 2\pi j) = P(\xi)$ for all j in Z^n . Here $S(R^n)$ is the Schwartz space of rapidly decreasing functions on R^n . It follows from (7), the periodicity of P , and the fact that \hat{f} is square integrable that

$$Q(\xi) = (2\pi)^{n/2} P(\xi) \left(\sum_{j \in Z^n} |\xi - 2\pi j|^{-4k} \right)^{1/2}$$

is locally square integrable. Hence we may write

$$\hat{f}(\xi) = Q(\xi) \hat{\Phi}_k(\xi)$$

or

$$f(x) = \sum_{j \in \mathbb{Z}^n} a_j \Phi_k(x - j)$$

where

$$Q(\xi) = \sum_{j \in \mathbb{Z}^n} a_j e^{-i(j, \xi)} \text{ and } \sum_{j \in \mathbb{Z}^n} |a_j|^2 < \infty.$$

As a result we may conclude that (6) is a complete orthonormal set in V_0 and the family $\mathcal{V}(k)$ satisfies property (v).

The fact that the family $\mathcal{V}(k)$ enjoys properties (ii) and (iii) is an immediate consequence of the definition. Hence, to conclude that $\mathcal{V}(k)$ is a dyadic multiscale analysis, it remains to show that it satisfies property (i).

First, to see that $\bigcap_{m=-\infty}^{\infty} V_m(k) = \{0\}$ observe that if f is in $\bigcap_{m=-\infty}^{\infty} V_m(k)$ then f must be a k -harmonic polynomial in $L^2(\mathbb{R}^n)$ and consequently $f \equiv 0$.

To conclude that $\bigcup_{m=-\infty}^{\infty} V_m(k)$ is dense in $L^2(\mathbb{R}^n)$ we argue as follows: Let f be any element in $L^2(\mathbb{R}^n)$ such that \hat{f} is bounded and has compact support. Define s_m to be the k -harmonic spline whose Fourier transform is

$$\hat{s}_m(\xi) = P_m(\xi) |\xi|^{-2k}$$

where P_m is the $2^{m+1}\pi$ periodization of $|\xi|^{2k} \hat{f}(\xi)$, namely,

$$P_m(\xi) = \sum_{j \in \mathbb{Z}^n} |\xi - 2^{m+1}\pi j|^{2k} \hat{f}(\xi - 2^{m+1}\pi j).$$

Clearly s_m is in $V_m(k)$. Now for m sufficiently large we may write

$$\begin{aligned} \|f - s_m\|_{L^2(\mathbb{R}^n)}^2 &= \int_{\mathbb{R}^n} |\hat{f}(\xi) - \hat{s}_m(\xi)|^2 d\xi \\ &= \int_{|\xi| \geq 2^m \pi} |\hat{s}_m(\xi)|^2 d\xi \leq C \int_{|\xi| \geq 2^m \pi} |\xi|^{-4k} d\xi \end{aligned}$$

where C is a constant independent of m . Finally, since the last integral in the above string clearly goes to 0 as m goes to infinity and the class of such f 's is dense in $L^2(\mathbb{R}^n)$, we may conclude that the desired result holds.

We summarize these observations as follows:

Proposition 2 *If $4k > n$ then the increasing sequence of subspaces $\mathcal{V}(k)$ defined by (2) is a dyadic multiscale analysis of $L^2(\mathbb{R}^n)$. Furthermore, if Φ_k is defined by (5) then the collection (6) is a complete orthonormal basis for $V_0(k)$.*

Note that if $2k \leq n$ then the non-zero elements of $V_0(k)$ are not locally bounded. On the other hand if $2k > n$ then the elements f of $V_0(k)$ are continuous and enjoy the representation

$$(8) \quad f(x) = \sum_{j \in \mathbb{Z}^n} f(j) L_k(x - j)$$

where L_k is defined by the formula for its Fourier transform

$$(9) \quad \hat{L}_k(\xi) = (2\pi)^{-n/2} \frac{|\xi|^{-2k}}{\sum_{j \in \mathbb{Z}^n} |\xi - 2\pi j|^{-2k}}.$$

and the series converges uniformly for x on compact subsets of R^n . For each such f we may write

$$\begin{aligned} \|f\|_{L^2(R^n)}^2 &= \int_{R^n} |(2\pi)^{n/2} \varphi \hat{f}(\xi) \hat{L}_k(\xi)|^2 d\xi \\ &= \int_{Q^n} |\varphi \hat{f}(\xi)|^2 \frac{\sum_{j \in \mathbb{Z}^n} |\xi - 2\pi j|^{-4k}}{\left(\sum_{j \in \mathbb{Z}^n} |\xi - 2\pi j|^{-2k}\right)^2} d\xi \end{aligned}$$

where

$$(2\pi)^{n/2} \varphi \hat{f}(\xi) = \sum_{j \in \mathbb{Z}^n} f(j) e^{-i(j, \xi)}.$$

Since there are constants $0 < C_1 \leq C_2$, independent of ξ such that

$$C_1 \leq \frac{\sum_{j \in \mathbb{Z}^n} |\xi - 2\pi j|^{-4k}}{\left(\sum_{j \in \mathbb{Z}^n} |\xi - 2\pi j|^{-2k}\right)^2} \leq C_2$$

we may conclude from the last formula for $\|f\|_{L^2(R^n)}^2$ that

$$C_1 \sum_{j \in \mathbb{Z}^n} |f(j)|^2 \leq \|f\|_{L^2(R^n)}^2 \leq C_2 \sum_{j \in \mathbb{Z}^n} |f(j)|^2$$

whenever f is in $V_0(k)$ and $2k > n$.

In other words we have the following:

Proposition 3 *If $2k > n$ then the increasing sequence of subspaces $\mathcal{V}(k)$ defined by (2) is a dyadic multiscale analysis of $L^2(R^n)$ which is composed of spline functions in the sense of Y. Meyer. Furthermore, in this case the complete orthonormal basis for $\mathcal{V}(k)$ defined by (5) and (6) enjoys exponential decay, namely,*

$$|\Phi_k(x)| \leq C e^{-\epsilon |x|}$$

and where C and ϵ are positive constants which depend only on k and n .

3 Entire functions of exponential type

Suppose f is a tempered distribution on R^n whose Fourier transform \hat{f} is supported in

$$Q^n = \{\xi = (\xi_1, \dots, \xi_n) : -\pi \leq \xi_j \leq \pi, j = 1, \dots, n\}.$$

It is a consequence of a well known generalization of a Paley-Wiener theorem that such an f is analytic, is of polynomial growth on R^n , and its holomorphic extension to complex n -space is an entire function of exponential type. The class of all such f 's will be denoted by $E_\pi(R^n)$.

In particular, given such an f the sequence of values

$$(10) \quad \{f(\mathbf{j}) : \mathbf{j} \in Z^n\}$$

is well defined. If, in addition, f is in $L^2(R^n)$ then it is well known that f can be recovered from these values via the formula

$$(11) \quad f(x) = \sum_{\mathbf{j} \in Z^n} f(\mathbf{j}) \operatorname{sinc}(x - \mathbf{j})$$

where

$$\operatorname{sinc}(x) = \operatorname{sinc}(x_1, \dots, x_n) = \prod_{j=1}^n \frac{\sin(\pi x_j)}{\pi x_j}.$$

Here the series converges uniformly and absolutely on compact subsets of R^n ; it also converges in $L^2(R^n)$. Similar results hold if the sequence of values (10) is in $\ell^p(Z^n)$ for some p , $1 < p < \infty$. Unfortunately, if the sequence of values (10) fails to be in $\ell^p(Z^n)$ for some p , $1 < p < \infty$, the series (11) may not make sense. On the other hand, since the sequence (10) is of polynomial growth the k -harmonic spline interpolant

$$(12) \quad S_k f(x) = \sum_{\mathbf{j} \in Z^n} f(\mathbf{j}) L_k(x - \mathbf{j})$$

is well defined for all f in $E_\pi(R^n)$ and all k such that $2k > n$. In view of the fact that

$$\lim_{k \rightarrow \infty} L_k(x) = \operatorname{sinc}(x)$$

it is natural to use (12) to approximate f with the expectation that

$$(13) \quad \lim_{k \rightarrow \infty} S_k f(x) = f(x).$$

Such a development is reminiscent of classical summability theory for Fourier series. In what follows we will show that (13) holds for a relatively wide class of f 's in $E_\pi(R^n)$.

To set up our result we will need several technical definitions. First, recall that if T is a distribution with compact support then the periodization of T , $\wp T$, can be defined by

$$\langle \wp T, \phi \rangle = \langle T, \wp \phi \rangle$$

where ϕ is any test function in the Schwartz class $\mathcal{S}(R^n)$,

$$\wp \phi(\xi) = \sum_{\mathbf{j} \in Z^n} \phi(\xi - 2\pi \mathbf{j}),$$

and $\langle T, \phi \rangle$ denotes the distribution T acting on the test function ϕ . Note that $\wp T$ is a well defined periodic tempered distribution.

Observe that if f is in $E_\pi(R^n)$ then

$$(14) \quad \widehat{S_k f}(\xi) = (2\pi)^{n/2} \wp f(\xi) \hat{L}_k(\xi)$$

where the multiplication on the right hand side makes sense because \hat{L}_k is smooth.

Define the functions λ and χ by the formulas

$$(15) \quad (2\pi)^{-n/2} \chi(\xi) = \lambda(\xi) = \lim_{k \rightarrow \infty} \hat{L}_k(\xi).$$

Note that $\chi(\xi)$ is 1 if ξ is in the interior of Q^n and 0 if ξ is in $R^n \setminus Q^n$. If ξ is in the boundary of Q^n then $\chi(\xi) = 2^{-m}$ if exactly m components of ξ have modulus π .

Finally, recall that T is said to be a distribution of order 0 in an open set Ω if

$$|\langle T, \phi \rangle| \leq C \|\phi\|_{L^\infty}$$

for all ϕ in $C_0^\infty(\Omega)$ where C is a constant independent of ϕ . We remind the reader that given such a T , by virtue of Riesz representation, there is a bounded Borel measure μ such that

$$\langle T, \phi \rangle = \int_{\Omega} \phi(\xi) d\mu(\xi).$$

Furthermore, if Ω_1 is an open set whose closure is in Ω and g is any bounded function which is infinitely differentiable on $R^n \setminus \Omega_1$ then the product $g(\xi)T(\xi)$ is well defined as a distribution on R^n . Namely, for any test function ψ this product is defined by

$$\begin{aligned} \langle gT, \psi \rangle &= \langle gT, \psi\phi \rangle + \langle gT, \psi(1-\phi) \rangle \\ &= \int_{\Omega} g(\xi)\psi(\xi)\phi(\xi)d\mu(\xi) + \langle T, g\psi(1-\phi) \rangle \end{aligned}$$

where ϕ is a function in $C_0^\infty(\Omega)$ such that $\phi(\xi) = 1$ for ξ in Ω_1 . Note that this definition is independent of the particular choice of ϕ and that the resulting product is tempered whenever T is tempered.

Proposition 4 Suppose f is in $E_\pi(R^n)$, \hat{f} is a distribution of order 0 in some neighborhood of the boundary of Q^n , and $\chi(\xi)\wp \hat{f}(\xi) = \hat{f}(\xi)$. Then

$$\lim_{k \rightarrow \infty} S_k f(x) = f(x)$$

uniformly on compact subsets of R^n .

Proof Choose ϵ , $0 < \epsilon < \pi$, such that \hat{f} is a distribution of order zero in

$$\Omega = \{\xi : \text{distance from } \xi \text{ to the boundary of } Q^n \text{ is } < \epsilon\}.$$

Set

$$\Omega_1 = \{\xi : \text{distance from } \xi \text{ to the boundary of } Q^n \text{ is } < \epsilon/2\}.$$

Let ϕ_1 be a function in $C_0^\infty(\Omega)$ such that $\phi_1(\xi) = 1$ for ξ in Ω_1 and let ϕ_0 and ϕ_2 be infinitely differentiable functions such that $\phi_0(\xi) + \phi_1(\xi) + \phi_2(\xi) = 1$ for all ξ in R^n , ϕ_0 has support in $Q^n \setminus \Omega_1$, and ϕ_2 has support in $R^n \setminus (Q^n \cup \Omega_1)$.

Now, if e_x denotes the exponential $e_x(\xi) = e^{i(x, \xi)}$ we may write

$$\begin{aligned} (16) \quad S_k f(x) - f(x) &= \langle \hat{L}_k \hat{f}, e_x \rangle - (2\pi)^{-n/2} \langle \hat{f}, e_x \rangle \\ &= \langle (\hat{L}_k - \lambda) \hat{f}, e_x \rangle = \sum_{j=0,1,2} \langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_j \rangle. \end{aligned}$$

Consider each term which comprises the sum in the last expression.

First

$$\begin{aligned} |\langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_0 \rangle| &= |\langle \hat{f}, (\hat{L}_k - \lambda) e_x \phi_0 \rangle| \\ &\leq C \sum_{|\nu| \leq N} \sup_{\xi \in Q^n \setminus \Omega_1} |D^\nu((\hat{L}_k(\xi) - \lambda(\xi)) e_x(\xi) \phi_0(\xi))| \end{aligned}$$

since f is a distribution of order no greater than N . Because for each ν $|D^\nu(\hat{L}_k(\xi) - \lambda(\xi))|$ goes to zero as $k \rightarrow \infty$ uniformly for ξ in $Q^n \setminus \Omega_1$ and $|D^\nu e_x(\xi)| \leq C_\nu |x|^{|\nu|}$ it follows from the last equality/inequality string that

$$(17) \quad \lim_{k \rightarrow \infty} |\langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_0 \rangle| = 0$$

uniformly for x in bounded subsets of R^n .

Next, there is a bounded Borel measure μ such that

$$\langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_1 \rangle = \int_{\Omega} (\hat{L}_k(\xi) - \lambda(\xi)) e_x(\xi) \phi_1(\xi) d\mu(\xi)$$

and, by virtue of the bounded convergence theorem, we may conclude that

$$(18) \quad \lim_{k \rightarrow \infty} |\langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_1 \rangle| = 0$$

uniformly in x .

Finally,

$$\begin{aligned} |\langle (\hat{L}_k - \lambda) \hat{f}, e_x \phi_2 \rangle| &= |\langle \hat{f}, \hat{f}[(\hat{L}_k - \lambda) e_x \phi_2] \rangle| \\ &\leq C \sum_{j \in \mathbb{Z}^n} \sum_{|\nu| \leq N} \sup_{\xi \in Q_j^n} |D^\nu((\hat{L}_k(\xi) - \lambda(\xi)) e_x(\xi) \phi_2(\xi))| \end{aligned}$$

where $Q_j^n = \{\xi : \xi - 2\pi j \text{ is in } Q^n\}$. Note that the term corresponding to $j = 0$ is 0 because of the factor ϕ_2 and that the series converges whenever $2k > n$. Estimating the size of each term in the last sum following essentially the same reasoning as used in the case of the term involving ϕ_0 allows us to conclude that

$$(19) \quad \lim_{k \rightarrow \infty} |\langle (\hat{L}_k - \lambda) \phi f, e_x \phi_2 \rangle| = 0$$

uniformly for x in bounded subsets of R^n .

Formulas (17), (18), (19), and (16) imply the desired result. ■

Corollary 1 *Suppose f is in $E_\pi(R^n)$, \hat{f} is a distribution of order 0 in some neighborhood of the boundary of Q^n . Then*

$$\lim_{k \rightarrow \infty} S_k f(x) = f(x) + g(x)$$

uniformly on compact subsets of R^n where

$$\hat{g}(\xi) = \chi(\xi) \phi \hat{f}(\xi) - \hat{f}(\xi)$$

is a distribution with support in the boundary of Q^n

Examples show that the condition that \hat{f} be a measure in a neighborhood of the boundary of Q^n is not necessary. However other examples show that without some restriction on the behavior of \hat{f} near the boundary of Q^n the sequence $S_k f$ may fail to converge.

Many subclasses of distributions in $E_\pi(R^n)$ satisfy the hypothesis of Proposition 4. Consider the following transparent examples:

- f is in $L^p(R^n)$ for some p , $1 \leq p \leq 2$.
- $\Lambda^\alpha f$ is in $L^p(R^n)$ for some p , $1 \leq p \leq 2$ and some positive number α . Here $\widehat{\Lambda^\alpha f}(\xi) = |\xi|^\alpha \hat{f}(\xi)$. This includes various classes of distributions which themselves may not be in $L^p(R^n)$ but whose derivatives are.

Such examples include many of the classes considered by Schoenberg in the univariate case.

4 Remarks

Various forms and features of multiscale analysis have been known and used in harmonic analysis for some time. We cannot go into details here. The formal definitions used here were adopted from [9]. The current popularity of multiscale analysis is due to primarily to the efforts of R. R. Coifman and Y. Meyer and their collaborators. An introduction to the subject together with further developments and more references may be found in [5, 6, 9].

Orthogonal bases composed of spline functions (piecewise polynomials) in the univariate case have been considered quite early, for example see [4, 7]. However the introduction of such bases which are translates and dilates of one function (wavelet) is more recent; it can be found in [21].

In the case $n = 2$ the wavelets corresponding to $\mathcal{V}(k)$ can be easily constructed using the recipe given in [9]. Their construction in the general case is not so clear.

The distributional variant of the Paley-Wiener Theorem alluded to in the introduction to Section 3 can be found in [8]. For basic facts concerning sampling and the n -variate sinc series see [16].

At this workshop Sherman Riemenschneider pointed out that certain variants of Proposition 4 were known in the univariate case, see [17]. When f is in $E_\pi(R^n) \cap L^2(R^n)$ then (12) also holds in the $L^2(R^n)$ sense. Similar results hold for other subclasses of $E_\pi(R^n)$. For examples in the univariate case see [15, 19]; we are currently preparing various analogues in the n -variate case.

The material in [1, 2, 3, 9, 10, 11, 18] indicates that box splines also generate interesting examples of multiscale analyses. However, summability theory for regularly sampled functions in $E_\pi(R^n)$ using box splines instead of k -harmonic splines seems to be more difficult, see [1, 18] and the pertinent references cited there.

References

- [1] C. deBoor, K. Hollig, and S. D. Riemenschneider, Bivariate cardinal interpolation by splines on a three-direction mesh, *Ill. J. Math.*, 29 (1985), 533-566.
- [2] C. K. Chui, *Multivariate Splines*, CBMS Vol. 54, SIAM, Philadelphia, 1988.
- [3] C. K. Chui, K. Jetter, and J. D. Ward, Cardinal interpolation by multivariate splines, *Math Comp*, 48 (1987), 711-724.
- [4] Z. Ciesielski, P. Simon, and P. Sjölin, Equivalence of Haar and Franklin bases in L^p spaces, *Studia Math.*, 60 (1976), 195-211.
- [5] J. M. Combs, A. Grossman, and Ph. Tchamitchian, eds., *Wavelets*, Proceedings of the International Conference in Marseille, 1987, Springer-Verlag, Berlin, 1989.
- [6] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, 41 (1988), 909-996.
- [7] P. Franklin, A set of continuous orthogonal functions, *Math. Ann.*, 100 (1928), 522-529.

- [8] L. Hormander, *Linear Partial Differential Operators*, Springer-Verlag, New York, 1969.
- [9] S. Jaffard, P. G. Lemarier, S. Mallat, and Y. Meyer, Multiscale Analysis, manuscript, 24pp.
- [10] K. Jetter, A short survey on cardinal interpolation by box splines, in *Topics in Multivariate Approximation*, C. K. Chui, L. L. Schumaker, and F. I. Utreras, eds., Academic Press, New York, 1987, 125-139.
- [11] K. Jetter, Cardinal interpolation, submodules, and the 4-direction mesh, *Constr. Approx.*, 3 (1987), 169-188.
- [12] W. R. Madych, Cardinal interpolation with polyharmonic splines, to appear in *Multivariate Approximation IV*, W. Schempp and K. Zeller, eds., 1989.
- [13] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines, to appear in *J. Approx. Theory*.
- [14] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines: a minimization property, technical report, BRC/MATH-TR-89-1.
- [15] M. J. Marsden, F. B. Richards, and S. D. Riemenschneider, Cardinal spline interpolation operators on l^p data, *Indiana Univ. Math Journal*, 24 (1975), 677-689.
- [16] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons, Stuttgart, 1986.
- [17] S. D. Riemenschneider, Convergence of interpolating cardinal splines: power growth, *Israel J. Math.*, 23 (1976), 339-346.
- [18] S. D. Riemenschneider, Multivariate cardinal interpolation, in *Approximation Theory VI*, C. K. Chui, L. L. Schumaker, and J.D. Ward, eds., to appear.
- [19] I. J. Schoenberg, *Cardinal Spline Interpolation*, CBMS Vol. 12, SIAM, Philadelphia, 1973.
- [20] I. J. Schoenberg, On the remainders and convergence of cardinal spline interpolation for almost periodic functions, *Studies in Spline Functions and Approximation Theory*, S. Karlin *et al.* eds., Academic, New York, 1976.

- [21] J-O. Strömberg, A modified Franklin system and higher order spline systems on R^n as unconditional bases for Hardy spaces, in *Conference on Harmonic Analysis in Honor of Antoni Zygmund*, W. Beckner, A. P. Calderon, R. Fefferman, and P. W. Jones, eds., Wadsworth, 1983, 475-494.

Summability and Approximate Reconstruction from Radon Transform Data

W. R. Madych*

Abstract

We derive various reconstruction formulas using complete and incomplete Radon transform data. In particular, the limited angle problem is treated. These formulas are of convolution type and are reminiscent of summability theory for Fourier series. In cases where exact reconstruction is not possible, formulas for approximate reconstruction are given together with bounds on the error which are asymptotically optimal.

1 Introduction

Recall that the Radon transform of a scalar valued function f on R^n is defined by

$$(1) \quad \mathcal{R}f(u, t) = \int_{u^\perp} f(tu + y) dy$$

where u is an element of the unit sphere S^{n-1} in R^n , t is a real number, and $u^\perp = \{y \in R^n : \langle y, u \rangle = 0\}$ is the hyperplane through the origin which is perpendicular to u . Thus $\mathcal{R}f(u, t)$ is the integral of f over the hyperplane $\{y \in R^n : \langle y, u \rangle = t\}$ which is perpendicular to u and intersects tu . The mapping $f \rightarrow \mathcal{R}f$ maps functions f defined on R^n to functions $\mathcal{R}f$ defined on $S^{n-1} \times R$.

In what follows we will often consider $\mathcal{R}f(u, t)$ as a family of functions of the variable t parametrized by u . In such cases we will use the notation

$$(2) \quad f_u(t) = \mathcal{R}f(u, t).$$

Thus f_u is the function of the real variable t defined by (2).

A problem associated with \mathcal{R} which has attracted wide attention concerns the recovery of the phantom f from full or partial knowledge of $\mathcal{R}f$. In

*Department of Mathematics, University of Connecticut, Storrs, CT 06269. Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-89-0524

this report we address certain aspects of this problem. Our approach is motivated by the classical theory of Fourier series and the constructive theory of functions. The main ideas may be described as follows:

For certain classes of convolution kernels K , the convolution $K * f$ can be conveniently expressed in terms of $\mathcal{R}f$. Such K can be chosen to be approximates of the identity. Doing so results in approximations of f and, in the limit, leads to various reconstruction formulas and algorithms. This is reminiscent of summability theory for Fourier series.

In the case of incomplete data the class of kernels K for which $K * f$ may be conveniently expressed is considerably more narrow than in the case of full data $\mathcal{R}f$. Here we consider data of the form $\{\mathcal{R}f(u, t) : -\infty < t < \infty, u \in \mathcal{A}\}$ where \mathcal{A} is a proper subset of S^{n-1} . For certain subsets \mathcal{A} the problem of reconstruction from such data is sometimes referred to as the limited angle problem. We show that it is possible to express $K * f$ in terms of this data for a sufficiently wide class of kernels K to obtain various summability theorems and approximate reconstruction formulas which result in the "correct" degree of approximation.

1.1 Contents

This paper is organized as follows:

Section 2 contains a description of the basic results in the case of full data. Various examples are contained in Section 3. Certain technical results needed for the development of the partial data case are also contained in Section 3.

Section 4 is perhaps the most interesting; it contains a description of a class of reconstruction and approximate reconstruction formulas for f in terms of $\{f_u : u \in \mathcal{A}\}$ where \mathcal{A} is a proper subset of S^{n-1} . Error bounds are also included.

Various comments which are not particularly germane to the development but are pertinent to the subject matter are given in Section 5.

The introduction to each section contains a rough indication of its contents.

1.2 Notation

We now briefly digress to list some of the conventions which are used here:

The Fourier transform \hat{f} of a function f on R^n is defined by

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int_{R^n} e^{-i\langle \xi, x \rangle} f(x) dx$$

whenever it makes sense and distributionally otherwise. Basic facts concerning Fourier transforms and distributions will be used without further elaboration in what follows; for example, see [10]. The term support is also

used in the distributional sense; in particular the support of a measurable function f is a well defined closed set. $L^p(R^n)$ is the usual Lebesgue class of measurable functions on R^n , see [10]. If Ω is a subset of R^n then $L^p(\Omega)$ is the L^p closure of the subspace of those functions in $L^p(R^n)$ whose support is contained in Ω .

The symbol c is used to denote generic constants. Its value depends on the context.

2 Full data

As mentioned in the introduction, in this section we indicate how to obtain general formulas for approximations of f in terms of $\mathcal{R}f$ which are analogous to the classical summability formulas associated with, for example, the Fejer or Abel kernels in the theory of Fourier series. These formulas are based on certain elementary observations.

2.1 Summability Formulas

A function H on R^n is said to be a *ridge function* if $H(x) = h(\langle x, u \rangle)$ for some unit vector u and some univariate function h . If H is such a ridge function then convolving it with another function f on R^n and expressing the integral in the $\{u, u^\perp\}$ coordinate system results in

$$\begin{aligned} \int_{R^n} H(x-y)f(y)dy &= \int_{-\infty}^{\infty} \left(\int_{u^\perp} h(\langle x, u \rangle - t)f(tu + z)dz \right) dt \\ &= \int_{-\infty}^{\infty} h(\langle x, u \rangle - t)f_u(t)dt \end{aligned}$$

or, in more compact and suggestive notation,

$$(3) \quad H * f(x) = h * f_u(\langle x, u \rangle).$$

In other words, the multivariate convolution of H and f evaluated at x is equal to the univariate convolution of h and f_u evaluated at $\langle x, u \rangle$. We are assuming of course that all the functions involved are sufficiently well behaved so that the integrals make sense.

Formula (3) is elementary and, as we have seen, easily verifiable. Nevertheless it is very useful; indeed, it is the basis of the summability formulas given below. For instance it should be quite easy to see that if a convolution kernel K is a sum of ridge functions then $K * f$ can be readily computed from knowledge of $\mathcal{R}f$. The definitions and formulas below are simply more precise versions of this observation.

A locally integrable function K on R^n is a *uniform sum of ridge functions* if there is an even locally integrable univariate function h such that

$$(4) \quad K(x) = \frac{1}{\sigma_n} \int_{S^{n-1}} h(\langle x, u \rangle) d\sigma(u)$$

for almost all x in R^n . Here $d\sigma(u)$ denotes the usual rotation invariant Lebesgue measure on the unit sphere S^{n-1} in R^n and σ_n denotes its total measure.

Observe that, as a consequence of the definition, if K is a uniform sum of ridge functions then it must be a radial function. In other words, $K(x) = k(|x|)$ where k is a univariate function.

The proposition below gives a formula for $K * f$ in terms of $\mathcal{R}f$. It is a transparent consequence of (3), (1), and an appropriate change in the order of integration which can be justified via Fubini's Theorem.

Proposition 1 *If a locally bounded function K is a uniform sum of ridge functions and f is an integrable function with compact support then*

$$(5) \quad K * f(x) = \frac{1}{\sigma_n} \int_{S^{n-1}} h * f_u(\langle x, u \rangle) d\sigma(u)$$

where h and K are related by (4).

The hypothesis on f and K in the above proposition can be relaxed in various ways. For example, if both K and h are integrable over R^n and R respectively then f need not have compact support. We will not go into further details here since the proposition is sufficiently general for our applications.

Formula (5) is very suggestive; it should allow us to obtain summability results analogous to the classical summability results alluded to in the beginning of this section. In view of this, the question concerning what radial functions are uniform sums of ridge functions is of some interest.

2.2 Uniform Sums of Ridge Functions

To obtain an answer to the question concerning what radial functions are uniform sums of ridge functions we begin by writing

$$(6) \quad K(x) = k(|x|)$$

to express the relationship between the radial function K and the corresponding univariate function k . Using this notation and the Funk-Hecke theorem, we may rewrite formula (4) as

$$(7) \quad k(r) = \frac{\sigma_{n-1}}{\sigma_n} \int_{-1}^1 h(rt)(1-t^2)^{(n-3)/2} dt$$

where $r = |x|$. This can also be derived more directly by using polar coordinates to express the integral on the right hand side of (4) as

$$\sigma_{n-1} \int_0^\pi h(r \cos \theta)(\sin \theta)^{n-2} d\theta$$

followed by the change of variables $t = \cos \theta$. Another useful variant of (4)

$$(8) \quad k(r) = \frac{2\sigma_{n-1}}{\sigma_n r^{n-2}} \int_0^r h(s)(r^2 - s^2)^{(n-3)/2} ds$$

follows from using the change of variables $s = rt$ and the fact that h is an even function.

It is clear from formula (8) that for given k the function h is unique. A representation of h in terms of k follows from (8) and the identities on pages 14 and 15 in [15]. An alternate approach results from the observation that the change of variables $\sigma = s^2$ and $\rho = r^2$ reduces (8) to a standard half line convolution equation which can be solved by routine methods, for example, Laplace transforms. In any case a representation of h in terms of k is given by

$$(9) \quad h(t) = c_n t \left(\frac{1}{2t} \frac{d}{dt} \right)^{n-1} \int_0^t (t^2 - s^2)^{(n-3)/2} s^{n-1} k(s) ds$$

which may also be expressed as

$$(10) \quad h(t) = c_n t \left(\frac{1}{2t} \frac{d}{dt} \right)^{n-1} t^{2n-3} \int_0^1 (1 - s^2)^{(n-3)/2} s^{n-1} k(ts) ds$$

where

$$c_n = \frac{2\sqrt{\pi}}{\Gamma((n-1)/2)\Gamma(n/2)}.$$

These observations can be summarized as follows:

Proposition 2 (i) If K is a uniform sum of ridge functions then K is radial and the h in representation (4) is unique. (ii) Conversely, if K is a radial function, $K(x) = k(|x|)$, and k is $n-1$ times continuously differentiable then K is a uniform sum of ridge functions. Furthermore, the h in representation (4) is unique and can be obtained from formula (9) or (10).

The conditions on k in the second half of the statement of the proposition can be somewhat relaxed. In the cases $n = 2$ and $n = 3$ the details are given below.

2.2.1 The Cases $n = 2$ and $n = 3$

In the case $n = 2$ equation (9) may be written more explicitly as

$$h(t) = \frac{d}{dt} \int_0^t (t^2 - s^2)^{-1/2} sk(s) ds.$$

Integration by parts followed by the indicated differentiation results in

$$(11) \quad h(t) = k(0) + t \int_0^t (t^2 - s^2)^{-1/2} dk(s)$$

where the integral is a standard Stieltjes integral.

The above calculations are valid whenever k is a function of bounded variation on every finite interval $[0, T]$, $0 < T < \infty$. Furthermore, if k satisfies this hypothesis, it follows from Fubini's theorem that

$$\int_0^T \int_0^t t (t^2 - s^2)^{-1/2} dk(s) dt = \int_0^T (T^2 - s^2)^{1/2} dk(s).$$

By inserting absolute values and replacing $=$ with \leq in the appropriate places in the last formula one can readily see that h is a locally integrable function and $\int_0^T |h(t)| dt$ is bounded by T times the sum of $|k(0)|$ and the total variation of k over the interval $[0, T]$.

These observations can be summarized by the following addendum to part (ii) of Proposition 2.

Proposition 3 *Suppose K is a radial function on R^2 , $K(x) = k(|x|)$, and k is a function of bounded variation on the intervals $[0, T]$, $0 < T < \infty$. Then K is a uniform sum of ridge functions and the h in representation (4) can be obtained from formula (11).*

In the case $n = 3$ equation (9) reduces to

$$(12) \quad h(t) = k(t) + tk'(t)$$

where k' denotes the derivative of k . From this we may easily conclude the following:

Proposition 4 *Suppose K is a radial function on R^3 such that $K(x) = k(|x|)$ and k is absolutely continuous. Then K is a uniform sum of ridge functions and the h in representation (4) can be obtained from formula (12).*

2.3 Summability and reconstruction

The above results allow for a very large class of summability and reconstruction formulas.

To see this, suppose K is an integrable function on R^n , that is it satisfies

$$(13) \quad \int_{R^n} |K(x)| dx < \infty.$$

Normalize K so that

$$(14) \quad \int_{R^n} K(x) dx = 1$$

and for positive ϵ write

$$(15) \quad K_\epsilon(x) = \epsilon^{-n} K(x/\epsilon).$$

Recall that if f is in $L^p(R^n)$ for some p which satisfies $1 \leq p < \infty$, then $K_\epsilon * f$ converges to f in $L^p(R^n)$ as ϵ goes to zero. Furthermore if K also satisfies

$$(16) \quad |K(x)| \leq \Phi(|x|) \text{ where } \Phi \text{ is a nonincreasing function} \\ \text{on the interval } 0 < t < \infty \text{ and } \int_0^\infty \Phi(r)r^{n-1}dr < \infty$$

then $K_\epsilon * f(x)$ converges to $f(x)$ almost everywhere. Detailed proofs of these facts can be found in [52]. Estimates on the rates of convergence of $K_\epsilon * f$ to f can be had in term of various smoothness conditions on f and certain moment conditions on K . If such a K is a uniform sum of ridge functions then by virtue of Proposition 1 we can easily obtain the type of results alluded to above.

To state some of this more precisely, suppose K is a uniform sum of ridge functions. Then

$$(17) \quad K_\epsilon(x) = \frac{1}{\sigma_n} \int_{S^{n-1}} h_\epsilon(\langle x, u \rangle) d\sigma(u)$$

where $h_\epsilon(t) = \epsilon^{-n} h(t/\epsilon)$ and h is the function in representation (4) of K .

Proposition 5 *Suppose K is an integrable kernel which is a uniform sum of ridge functions and which satisfies (13), (14), and (16). Define the functions K_ϵ and h_ϵ via (15) and (17) and assume that h_ϵ is integrable. Suppose f is an integrable function. Then*

$$(18) \quad \lim_{\epsilon \rightarrow \infty} \frac{1}{\sigma_n} \int_{S^{n-1}} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u) = f(x)$$

for almost all x in R^n . If, in addition, f is in $L^p(R^n)$ for some p satisfying $1 \leq p < \infty$ then

$$\lim_{\epsilon \rightarrow \infty} \int_{R^n} |f(x) - \frac{1}{\sigma_n} \int_{S^{n-1}} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u)|^p dx = 0.$$

Furthermore, if f is uniformly continuous, (18) holds uniformly in x for $x \in R^n$.

As a specific example of K_ϵ satisfying the conditions of the above proposition consider the Poisson kernel

$$K_\epsilon(x) = \frac{c_n \epsilon}{(\epsilon^2 + |x|^2)^{(n+1)/2}}$$

where c_n is constant. That this kernel is a uniform sum of ridge functions follows from Proposition 2. The formula for the corresponding ridge function is

$$(19) \quad h_\epsilon(t) = c_n \frac{T_n(\epsilon/\sqrt{\epsilon^2 + t^2})}{(\epsilon^2 + t^2)^{n/2}}$$

where T_n is the classical Tchebichef polynomial of the first kind of degree n , see Subsection 3.2.1.

Corollary 1 Suppose h_ϵ is defined by (19) and suppose f is an integrable function. Then

$$(20) \quad f(x) = \lim_{\epsilon \rightarrow \infty} \frac{1}{\sigma_n} \int_{S^{n-1}} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u)$$

for almost all x in R^n . If, in addition, f is in $L^p(R^n)$ for some p satisfying $1 \leq p < \infty$ then

$$\lim_{\epsilon \rightarrow \infty} \int_{R^n} |f(x) - \frac{1}{\sigma_n} \int_{S^{n-1}} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u)|^p dx = 0.$$

Furthermore, if f is bounded and uniformly continuous, (20) holds uniformly in x .

These results can be used to obtain elementary derivation of various variants of the usual type of "inversion formulas". For example, in the case $n = 2$ formula (19) gives

$$(21) \quad h_\epsilon(t) = \frac{1}{2\pi} \frac{\epsilon^2 - t^2}{(\epsilon^2 + t^2)^2}.$$

Since $\int_{-\infty}^{\infty} h_\epsilon(t) dt = 0$ we may write

$$(22) \quad \begin{aligned} \frac{1}{2\pi} \int_{S^1} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u) &= \\ &= \frac{1}{2\pi} \int_{S^1} \int_{-\infty}^{\infty} h_\epsilon(t) \{f_u(\langle x, u \rangle - t) - f_u(\langle x, u \rangle)\} dt d\sigma(u) \end{aligned}$$

or, using the fact that h_ϵ is even,

$$(23) \quad \begin{aligned} \frac{1}{2\pi} \int_{S^1} h_\epsilon * f_u(\langle x, u \rangle) d\sigma(u) &= \\ &= \frac{1}{4\pi} \int_{S^1} \int_{-\infty}^{\infty} h_\epsilon(t) \{f_u(\langle x, u \rangle + t) - 2f_u(\langle x, u \rangle) + f_u(\langle x, u \rangle - t)\} dt d\sigma(u). \end{aligned}$$

Now, assuming that f_u is sufficiently smooth, the right hand side of (22) may be expressed as

$$\frac{-1}{2\pi} \int_{S^1} \int_0^1 \int_{-\infty}^{\infty} t h_{s\epsilon}(t) f'_u(\langle x, u \rangle - t) dt ds d\sigma(u)$$

where f'_u denotes the derivative of f_u . Since $\lim_{\epsilon \rightarrow 0} t h_{s\epsilon}(t) = -1/(2\pi t)$, it can be anticipated that if ϵ goes to 0 then the innermost integral in the last expression will converge to the Hilbert transform of f'_u ; for examples of such arguments see [53]. This will result in an integrand which is independent of the s and so this variable can be integrated out.

Letting ϵ go to zero in (23) also results in an expression which is independent of ϵ . We summarize these elementary calculations as follows:

Corollary 2 *If f is a function on R^2 which is twice continuously differentiable and has compact support then*

$$(24) \quad f(x) = \frac{1}{4\pi^2} \int_{S^1} \int_{-\infty}^{\infty} \frac{f'_u(\langle x, u \rangle - t)}{t} dt d\sigma(u)$$

and

$$f(x) = \frac{-1}{8\pi^2} \int_{S^1} \int_{-\infty}^{\infty} \frac{f_u(\langle x, u \rangle + t) - 2f_u(\langle x, u \rangle) + f_u(\langle x, u \rangle - t)}{t^2} dt d\sigma(u)$$

for all x . The integral in the t variable in (24) is to be interpreted in the principal value sense and f'_u denotes the derivative of f_u .

Formula (24) is one of the standard variants of the Radon inversion formula and may be found, for instance, in [41]. We have not attempted to give the most general hypothesis on f in the last corollary; by going through the above calculations the reader should have no difficulty in relaxing the conditions on f somewhat. It may be of some interest to obtain the weakest possible conditions on f for which the formulas in the above corollary are valid almost everywhere.

3 Examples and miscellaneous results

The first three subsections below contain examples of ridge function representations and alternate ways of computing such representations. The final three subsections contain material which will be used in our treatment of partial data.

3.1 Ridge function representations for certain bivariate kernels

Because it models certain non-destructive imaging schemes, the bivariate case is particularly important in practical applications. For example see [8, 16, 41]. In this subsection we give several examples in the case $n = 2$ of radial approximations of the identity and their ridge function representations.

3.1.1 Averaging kernels

If K is $1/\pi$ times the indicator function of the unit disk centered at the origin then $K_\epsilon * f(x)$ is the average of f over the disk centered at x . The corresponding ridge function h_ϵ can be easily derived from (11). Thus if

$$K_\epsilon(x) = \begin{cases} 1/(\pi\epsilon^2) & \text{if } |x| \leq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

then the corresponding ridge function in representation (4) is given by

$$h_\epsilon(t) = \begin{cases} 1/(\pi\epsilon^2) & \text{if } |t| \leq \epsilon \\ \{1/(\pi\epsilon^2)\} \{1 - |t|/\sqrt{t^2 - \epsilon^2}\} & \text{otherwise.} \end{cases}$$

3.1.2 Multiples of $(1 + |x|^2)^\alpha$.

If $K(x) = (1 + |x|^2)^\alpha$ then its corresponding ridge function $h(t)$ can be obtained from (11) quite conveniently. Indeed,

$$h(t) = 1 + 2\alpha t \int_0^t (1 + s^2)^{\alpha-1} (t^2 - s^2)^{-1/2} s ds$$

which, after appropriate change of variables, can be expressed as

$$(25) \quad h(t) = 1 + 2\alpha t (1 + t^2)^{\alpha-1/2} \int_0^{t/\sqrt{1+t^2}} (1 - s^2)^{\alpha-1} ds.$$

In the cases when α is an integer multiple of $1/2$ the integral in (25) can be easily evaluated and expressed in terms of elementary functions. For example, in the case $\alpha = -3/2$ formula (25) reduces to

$$h(t) = \frac{1 - t^2}{(1 + t^2)^2}.$$

Thus if K_ϵ is the so-called Poisson kernel

$$K_\epsilon(x) = \frac{1}{2\pi} \frac{\epsilon}{(\epsilon^2 + |x|^2)^{3/2}}$$

then the corresponding ridge function in representation (4) is given by

$$h_\epsilon(t) = \frac{1}{2\pi} \frac{\epsilon^2 - t^2}{(\epsilon^2 + t^2)^2}.$$

3.2 Fourier transforms

The notion of uniform sum of ridge functions expressed by (4) can be extended to tempered distributions as follows: Recall that the usual way of defining linear transformations on distributions is via duality and observe that the formal dual of the mapping

$$\phi(t) \longrightarrow \int_{S^{n-1}} \phi(\langle x, u \rangle) d\sigma(u)$$

which maps sufficiently well behaved univariate functions into n -variate functions is the mapping

$$\phi(x) \longrightarrow \int_{S^{n-1}} \mathcal{R}\phi(u, t) d\sigma(u).$$

Hence it is quite natural to define a distribution K in $\mathcal{S}'(R^n)$ to be a uniform sum of ridge distributions if there is an even distribution h in $\mathcal{S}'(R)$ such that

$$(26) \quad \langle K, \phi \rangle = \frac{1}{\sigma_n} \int_{S^{n-1}} \langle h, \phi_u \rangle d\sigma(u)$$

for all ϕ in $S(R^n)$. Here $\langle \alpha, \beta \rangle$ denotes the action of the distribution α on the test function β . We also remind the reader that $\phi_u(t) = \mathcal{R}\phi(u, t)$.

Note that for fixed u in S^{n-1} the mapping $\phi \rightarrow \phi_u$ maps $S(R^n)$ continuously into $S(R)$. Thus the tempered distribution K is well defined by (26) for every even distribution h in $S'(R)$.

Having defined (4) distributionally, its Fourier transform is also well defined as follows: If K satisfies (26) then its Fourier transform \hat{K} satisfies

$$(27) \quad \langle \hat{K}, \phi \rangle = \frac{1}{\sigma_n} \int_{S^{n-1}} \langle \hat{h}, \phi^u \rangle d\sigma(u)$$

for all ϕ in $S(R^n)$. Here $\phi^u(t) = \phi(tu)$. Of course \hat{K} and \hat{h} are the Fourier transforms of K and h as distributions on R^n and R respectively. In the case when \hat{h} is sufficiently smooth, (27) may be expressed as

$$(28) \quad \hat{K}(\xi) = \frac{\hat{h}(|\xi|)}{\sigma_n |\xi|^{n-1}}.$$

3.2.1 The Poisson kernel

Formula (28) is useful in determining the ridge function representation of certain radial K 's. For example consider the Poisson kernel which is defined for positive ϵ by

$$(29) \quad K_\epsilon(x) = \frac{c_n \epsilon}{(\epsilon^2 + |x|^2)^{(n+1)/2}}$$

and whose Fourier transform is given by

$$(30) \quad \hat{K}_\epsilon(\xi) = (2\pi)^{-n/2} e^{-\epsilon|\xi|}.$$

Here c_n is a constant whose exact expression is

$$c_n = \pi^{(n+1)/2} \Gamma((n+1)/2).$$

The formula for the Fourier transform of the corresponding h_ϵ is easily obtained from (28) and (30); it is

$$(31) \quad \hat{h}_\epsilon(\tau) = (2\pi)^{-n/2} \sigma_n |\tau|^{n-1} e^{-\epsilon|\tau|}.$$

To obtain a formula for h_ϵ observe that (31) may be expressed as

$$\hat{h}_\epsilon(\tau) = (2\pi)^{-n/2} \sigma_n \left(-\frac{d}{d\epsilon} \right)^{n-1} e^{-\epsilon|\tau|}$$

and recall that $\exp(-\epsilon|\tau|)$ is the Fourier transform of a constant multiple of $\epsilon/(\epsilon^2 + t^2)$. Thus h_ϵ is a constant multiple of

$$\left(-\frac{d}{d\epsilon} \right)^{n-1} \frac{\epsilon}{(\epsilon^2 + t^2)} = c \frac{T_n(\epsilon/\sqrt{\epsilon^2 + t^2})}{(\epsilon^2 + t^2)^{n/2}}$$

where the equality follows from an induction argument, c is a constant, and T_n is the classical Tchebichef polynomial of the first kind of degree n , namely, $T_n(\cos \theta) = \cos n\theta$. This constant can be determined by evaluating both K_ϵ and h_ϵ at the origin which results in

$$(32) \quad h_\epsilon(t) = c_n \frac{T_n(\epsilon/\sqrt{\epsilon^2 + t^2})}{(\epsilon^2 + t^2)^{n/2}}$$

where c_n is the same constant as in (29).

3.3 Polynomials and ridge functions

Since it is quite transparent that

$$(33) \quad \int_{S^{n-1}} \langle x, u \rangle^{2m} d\sigma(u) = a(n, m) |x|^{2m}$$

where $a(n, m)$ is a positive constant, ridge function representations of radial polynomials is rather appealing. Thus if P is any radial polynomial

$$P(x) = \sum_{j=1}^M b_j |x|^{2j}$$

then

$$P(x) = \int_{S^{n-1}} h(\langle x, u \rangle) d\sigma(u)$$

where

$$h(t) = \sum_{j=1}^M \frac{b_j}{a(n, j)} t^{2j}.$$

The above formulas lead to several interesting relations between various classes of classical orthogonal polynomials. This can be realized by observing that

$$a(n, m) = \sigma_{n-1} \int_{-1}^1 t^{2m} (1 - t^2)^{(n-3)/2} dt$$

and, as in [30], using the representation of these polynomials in terms of the hypergeometric function. However a more intuitive approach can be outlined as follows:

Let $L^2(B^n, \lambda)$ be the Hilbert space consisting of those measurable functions f on the unit ball B^n for which

$$\int_{B^n} |f(x)|^2 (1 - |x|^2)^\lambda dx$$

is finite. This Hilbert space is equipped with the inner product

$$\langle f, g \rangle_\lambda = \int_{B^n} f(x) \overline{g(x)} (1 - |x|^2)^\lambda dx.$$

If \mathcal{P}_m denotes the set of n -variate polynomials of degree $\leq m$ then we define $L^2(B^n, \lambda, 0) = \mathcal{P}_0$ and for $m \geq 1$

$$L^2(B^n, \lambda, m) = \{P \in \mathcal{P}_m : \langle P, Q \rangle_\lambda = 0 \text{ for all } Q \in \mathcal{P}_{m-1}\}.$$

Finally, recall the classical Jacobi polynomials $P_m^{(\alpha, \beta)}(t)$, $m = 0, 1, 2, \dots$, and the classical Gegenbauer or ultraspherical polynomials $C_m^\lambda(t)$, $m = 0, 1, 2, \dots$; these are families of univariate polynomials which are orthogonal on the interval $-1 \leq t \leq 1$ with respect to the weights $(1+t)^\alpha(1-t)^\beta$ and $(1-t^2)^{\lambda-1/2}$ respectively; m denotes the degree of the polynomials, α , β , and λ are real parameters. The notation for these polynomials is standard, see [11].

Proposition 6 (i) A radial polynomial is in $L^2(B^n, \lambda, 2m)$ if and only if it is a constant multiple of $P_m^{(\lambda, (n-2)/2)}(2|x|^2 - 1)$. (ii) A ridge function is in $L^2(B^n, \lambda, m)$ if and only if it is a constant multiple of $C_m^{\lambda+n/2}(\langle x, u \rangle)$ for some unit vector u .

Proof If P is a radial polynomial in \mathcal{P}_{2m} then $P(x) = p(|x|^2)$ where p is a univariate polynomial of degree m and if Q is in \mathcal{P}_{2m-1} then

$$\langle P, Q \rangle_\lambda = \int_0^1 p(r^2)q(r^2)(1-r^2)^\lambda r^{n-1} dr$$

where

$$q(r^2) = \int_{S^{n-1}} \overline{Q(ru)} d\sigma(u)$$

and q is a univariate polynomials of degree $\leq m-1$. Hence

$$\langle P, Q \rangle_\lambda = \frac{1}{2} \int_0^1 p(t)q(t)(1-t)^\lambda t^{(n-2)/2} dt$$

and, since q may range over all polynomials of degree $\leq m-1$, (i) follows by virtue of the fact that constant multiples of $P_m^{(\lambda, (n-2)/2)}(2t-1)$ are the only polynomials of degree m which are orthogonal in this sense to all polynomials q of degree $\leq m-1$.

If P is a ridge polynomial in \mathcal{P}_m then $P(x) = p(\langle x, u \rangle)$ for some univariate polynomials of degree m and if Q is in \mathcal{P}_{m-1} then

$$\langle P, Q \rangle_\lambda = \int_{-1}^1 p(t)g(t)dt$$

where

$$g(t) = \int_{A(t)} \overline{Q(tu+y)} (1-t^2-|y|^2)^\lambda dy$$

and $A(t) = \{y \in u^\perp : |y|^2 \leq 1-t^2\}$. The the change of variable $y = (1-t^2)^{1/2}z$ allows us to express g as

$$g(t) = (1-t^2)^{\lambda+(n-1)/2} \int_{A(1)} \overline{Q(tu+(1-t^2)^{1/2}z)} (1-|z|^2)^\lambda dz.$$

Since $Q(tu + (1 - t^2)^{1/2}z) = \sum t^j (1 - t^2)^{|\nu|/2} z^\nu$ where the sum is taken over all integers j and multi-indices $\nu = (\nu_1, \dots, \nu_{n-1})$ such that $|\nu| + j = \nu_1 + \dots + \nu_{n-1} + j \leq m - 1$, the terms with odd $|\nu|$ integrate out to 0 and as a result $g(t) = q(t)(1 - t^2)^{\lambda+(n-1)/2}$ where q is a polynomial of degree $\leq m - 1$. Hence

$$\langle P, Q \rangle_\lambda = \int_{-1}^1 p(t)q(t)(1 - t)^{\lambda+n/2-1/2} dt$$

and, since q may range over all polynomials of degree $\leq m - 1$, (ii) follows by virtue of the fact that constant multiples of $C_m^{(\lambda+n/2)}(t)$ are the only polynomials of degree m which are orthogonal in this sense to all polynomials q of degree $\leq m - 1$. ■

Now, by virtue of (ii) of the above proposition $\int_{S^{n-1}} C_{2m}^{\lambda+n/2}(\langle x, u \rangle) d\sigma(u)$ is in $L^2(B^n, \lambda, 2m)$ and, since it is clearly radial, it must be a constant multiple of $P_m^{(\lambda, (n-2)/2)}(2|x|^2 - 1)$ view of (i) of this proposition. The constant can be easily evaluated by setting $x = 0$ in both expressions. As a result we may conclude the following:

Corollary 3 For $\lambda > -1$

$$(34) \quad P_m^{(\lambda, (n-2)/2)}(2|x|^2 - 1) = \frac{1}{\sigma_n} \int_{S^{n-1}} C_{2m}^{\lambda+n/2}(\langle x, u \rangle) d\sigma(u)$$

3.4 Moduli of smoothness, moment conditions, and degree of approximation

If f is in $L^p(R^n)$ for some value of p , $1 \leq p < \infty$, its L^p modulus of continuity $\omega_p(f; \delta)$, $\delta > 0$, is defined by

$$\omega_p(f; \delta) = \sup_{|y| \leq \delta} \left(\int_{R^n} |f(x - y) - f(x)|^p dx \right)^{1/p}.$$

Observe that this modulus is well defined and

$$(35) \quad \lim_{\delta \rightarrow 0} \omega_p(f; \delta) = 0$$

for all f in $L^p(R^n)$.

In the case $p = \infty$ this modulus reduces to the usual modulus of continuity, namely,

$$\omega_\infty(f; \delta) = \sup_{|x-y| \leq \delta} |f(x) - f(y)|.$$

Furthermore

$$\lim_{\delta \rightarrow 0} \omega_\infty(f; \delta) = 0$$

if and only if f is uniformly continuous.

We briefly summarize some of the elementary properties of these moduli. In what follows we take the parameter p and function f to be fixed and consider ω as a function of the parameter δ , $\delta > 0$; thus for simplicity we write $\omega_p(f; \delta) = \omega(\delta)$.

- The function $\omega(\delta)$ increases monotonically.
- For non-negative parameters δ_1 and δ_2 , $\omega(\delta_1 + \delta_2) \leq \omega(\delta_1) + \omega(\delta_2)$.
- For any positive integer n , $\omega(n\delta) \leq n\omega(\delta)$.
This is an immediate consequence of the previous item.
- If $\lambda > 0$, then $\omega(\lambda\delta) \leq (\lambda + 1)\omega(\delta)$.
To see this choose an integer n such that $n \leq \lambda \leq n + 1$ and write

$$\omega(\lambda\delta) \leq \omega((n + 1)\delta) \leq (n + 1)\omega(\delta) \leq (\lambda + 1)\omega(\delta).$$

- If $\delta_1 < \delta_2$ then

$$\frac{\omega(\delta_2)}{\delta_2} \leq 2 \frac{\omega(\delta_1)}{\delta_1}.$$

To see this simply write

$$\omega(\delta_2) = \omega\left(\frac{\delta_2}{\delta_1}\delta_1\right) \leq \left(\frac{\delta_2}{\delta_1} + 1\right)\omega(\delta_1) \leq 2\frac{\delta_2}{\delta_1}\omega(\delta_1).$$

These moduli provide a natural and useful measure of the degree of approximation of various processes. For example, suppose $K_\epsilon(x)$, $\epsilon > 0$, is a family of integrable convolution kernels which satisfy

$$(36) \quad \int_{R^n} K_\epsilon(x) dx = 1$$

$$(37) \quad \int_{R^n} |K_\epsilon(x)| dx \leq c_1$$

and

$$(38) \quad \int_{R^n} |x| |K_\epsilon(x)| dx \leq c_2 \epsilon$$

where c_1 and c_2 are constants independent of ϵ . Then it is not difficult to show the following:

Proposition 7 *If $K_\epsilon(x)$, $\epsilon > 0$, is a family of integrable convolution kernels which satisfy (36), (37), and (38) then for any f in $L^p(R^n)$, $1 \leq p \leq \infty$,*

$$\|K_\epsilon * f - f\|_{L^p(R^n)} \leq c\omega_p(f; \epsilon)$$

where c is a constant independent of f and ϵ .

Proof By virtue of (36) we may write

$$K_\epsilon * f(x) - f(x) = \int_{R^n} K_\epsilon(y) \{f(x-y) - f(x)\} dy.$$

Direct estimation in the cases $p = 1$ and ∞ or, in the cases $1 < p < \infty$, an application of the integral variant of Minkowski's inequality results in

$$\|K_\epsilon * f - f\|_{L^p(R^n)} \leq \int_{R^n} \omega_p(f; |y|) |K_\epsilon(y)| dy.$$

Finally, the estimate

$$\omega_p(f; |y|) = \omega_p\left(f; \frac{|y|}{\epsilon}\right) \leq \left(\frac{|y|}{\epsilon} + 1\right) \omega_p(f; \epsilon)$$

together with (37) and (38) imply the desired result. ■

Using the notation $\Delta_y f(x)$ to denote the difference $f(x-y) - f(x)$ and, if k is an integer ≥ 2 , $\Delta_y^k f(x) = \Delta_y(\Delta_y^{k-1} f)(x)$, we define the k -th order L^p modulus of smoothness $\omega_p^k(f; \delta)$, $\delta \geq 0$, via

$$\omega_p^k(f; \delta) = \sup_{|y| \leq \delta} \|\Delta_y^k f\|_{L^p(R^n)}.$$

Note that $\omega_p^1 = \omega_p$. In the case $k \geq 2$ the properties of ω_p^k are also well known; they are completely analogous to the case $k = 1$ with the obvious modifications, see [56]. Here we only point out the transparent facts that $\omega_p^m(f; \delta) \leq 2^{m-k} \omega_p^k(f; \delta)$ if $k \leq m$ and that $\omega_p^k(f; \delta)$ can go to zero as fast as $O(\delta^k)$ without being identically zero.

As in the case $k = 1$, these moduli provide a natural and useful measure of the degree of approximation for various processes. In the example where $K_\epsilon(x)$, $\epsilon > 0$, is a family of integrable convolution kernels, condition (38) ought to be replaced with

$$(39) \quad \int_{R^n} |x|^k |K_\epsilon(x)| dx \leq c_2 \epsilon^k$$

where c_2 is a constant independent of ϵ , in order to take advantage of the potentially faster decay as $\epsilon \rightarrow 0$. Then it is not difficult to show the following:

Proposition 8 Suppose $K_\epsilon(x)$, $\epsilon > 0$, is a family of integrable convolution kernels which satisfy (36), (37) and (39). Define

$$\tilde{K}_\epsilon^k(x) = \sum_{j=0}^{k-1} (-1)^{k-j-1} \frac{k!}{(k-j)!j!} (k-j)^{-n} K_\epsilon(x/(k-j)).$$

Then for any f in $L^p(R^n)$, $1 \leq p \leq \infty$,

$$\|\tilde{K}_\epsilon^k * f - f\|_{L^p(R^n)} \leq c \omega_p^k(f; \epsilon)$$

where c is a constant independent of f and ϵ .

Proof By virtue of (36) and the nature of \tilde{K}_ϵ^k we may write

$$\tilde{K}_\epsilon^k * f(x) - f(x) = \int_{R^n} K_\epsilon(y) \Delta_y^k f(x) dy.$$

The rest of the proof follows along the same lines as that of the previous proposition. ■

Before leaving this section, we remark that if $K(x)$ is an integrable function which satisfies (13), (14), and

$$(40) \quad \int_{R^n} |x|^k |K(x)| dx < \infty$$

then the family of convolution kernels $K_\epsilon(x) = \epsilon^{-n} K(x/\epsilon)$ satisfies (36), (37), and (39).

3.5 Polynomial convolution kernels

In many problems of practical and theoretical interest the phantom f is taken to have support in a fixed disk or ball of finite radius, say $1/2$, about the origin. In this case consideration of $K * f(x)$ for x in $B_{1/2}^n = \{x \in R^n : |x| \leq 1/2\}$ does not involve the behavior of K outside the ball of radius one about the origin. Similarly for such f and x the calculation of the quantities on the right hand side of formulas (5) does not involve the behavior of h outside the interval $-1 \leq t \leq 1$. In this case the criteria that K be a good approximation of the identity can be modified to

$$(41) \quad \int_{B^n} K(x) dx = 1$$

and

$$(42) \quad \int_{\delta < |x| \leq 1} |K(x)| dx < \epsilon$$

for sufficiently small δ and ϵ . A wide class of such kernels are given by certain radial polynomials. In view of the fact that ridge function representation of such polynomials is very appealing, the corresponding kernels should have significant applications in tomography.

There are many ways of producing radial polynomial kernels which behave like good approximations of the identity. A possible method of producing good kernels suitable for numerical work can be based on the Christoffel-Darboux formula for orthogonal polynomials.

For our purposes, summability and degree of approximation, we need polynomial kernels which, in addition to (41), enjoy appropriate analogues of (16), (37), and/or (39). These properties are easily verified for the families given below.

3.5.1 Kernels for summability

A family of radial polynomials which can be easily shown to possess an analogue of (16) is the following:

Consider the sequence of polynomials P_m , $m = 0, 1, 2, \dots$, defined by

$$(43) \quad P_m(x) = c(n, m)(1 - |x|^2)^m$$

where

$$(44) \quad \frac{1}{c(n, m)} = \int_{B^n} (1 - |x|^2)^m dx = \\ = \frac{\sigma_n}{2} \int_0^1 (1 - t)^m t^{n/2-1} dt = \frac{\sigma_n}{2} \frac{\Gamma(m+1)\Gamma(n/2)}{\Gamma(m+n/2+1)}$$

is a constant chosen so that

$$(45) \quad \int_{B^n} P_m(x) dx = 1$$

for all m . Observe that

$$(1 - |x|^2)^m \leq e^{-m|x|^2} \text{ whenever } |x| \leq 1$$

and

$$c(n, m) \leq c m^{n/2}$$

where c is a constant independent of m ; the last inequality follows easily from the definition of the gamma function Γ in the case of even n and by virtue of Stirling's formula in the case of odd n . These inequalities allow us to conclude that

$$(46) \quad P_m(x) \leq c m^{n/2} e^{-m|x|^2} \text{ whenever } |x| \leq 1$$

where c is a constant independent of m .

Observe that (46) is a special example of (16) in the case $\epsilon = m^{-1/2}$ with the restriction that $|x| \leq 1$. Hence we may conclude the following:

Proposition 9 *Suppose f is an integrable function with support in $B_{1/2}^n$, in other words $f \in L^1(B_{1/2}^n)$, and P_m , $m = 0, 1, 2, \dots$, is the sequence of polynomials defined by (43). Then*

$$(47) \quad \lim_{m \rightarrow \infty} P_m * f(x) = f(x)$$

for almost all x in $B_{1/2}^n$. If, in addition, f is in $L^p(B_{1/2}^n)$ for some p satisfying $1 \leq p < \infty$ then

$$\lim_{m \rightarrow \infty} \int_{B_{1/2}^n} |f(x) - P_m * f(x)|^p dx = 0.$$

Furthermore, if f is continuous, (47) holds uniformly in x for $x \in B_{1/2}^n$.

3.5.2 Degree of approximation

Unfortunately, the simple family of polynomial kernels defined by (43) is not sufficient to obtain the "correct" degree of approximation results. What is needed is a family which in addition to satisfying (45) also satisfies moment conditions of the form

$$\int_{B^n} |x|^k |P_m(x)| dx \leq \frac{c}{m^k}$$

where c is a constant independent of m . I know of no elementary formulas for such polynomials. Theoretically the most convenient and least technical construction seems to be the following:

The idea is, for each positive integer k , to construct a sequence of positive radial polynomials, P_m , $m = 1, 2, \dots$, where P_m is of degree no greater than $2m$, which satisfy

$$(48) \quad \int_{B^n} |x|^{2k} P_m(x) dx \leq \frac{c}{m^{2k}} \int_{B^n} P_m(x) dx$$

where c is a constant independent of m .

Using polar coordinates and the change of variable $t = r^2$ it is clear that it suffices to construct a sequence of positive univariate polynomials p_m such that

$$(49) \quad \int_0^1 p_m(t) t^{k+(n-2)/2} dt = \frac{c}{m^{2k}} \int_0^1 p_m(t) t^{(n-2)/2} dt$$

where p_m is of degree no greater than m and c is a constant independent of m .

The construction of such a sequence p_m is outlined as follows:

- Let N be the integer defined by

$$N = \begin{cases} (n-2)/2 & \text{if } n \text{ is even} \\ (n-1)/2 & \text{if } n \text{ is odd} \end{cases}$$

and let $\beta = (n-2)/2 - N$. Note that $\beta = 0$ or $-1/2$.

- Let ℓ be the integer defined by

$$\ell = k + N.$$

- Let M be the integer defined by

$$M = \begin{cases} (\ell + m + 2)/2 & \text{if } \ell + m \text{ is even} \\ (\ell + m + 1)/2 & \text{if } \ell + m \text{ is odd.} \end{cases}$$

- Let $0 < z_M(1) < z_M(2) < \dots < z_M(M) < 1$ be the zeros of $P_M^{(0,\beta)}(2t-1)$ where $P_M^{(0,\beta)}$ is the classical Jacoby polynomial of degree M and recall that for fixed j

$$(50) \quad z_M(j) \leq \frac{c}{M^2}$$

where c is a constant independent of M . See [55].

- Define

$$(51) \quad p_m(t) = \prod_{j=\ell+1}^M (t - z_M(j))^2$$

and observe that $p_m(t) t^N$ and $p_m(t) t^{k+N}$ are polynomials of degree no greater than $2M - 1$ so that Gaussian quadrature with knots at $z_M(1), z_M(2), \dots, z_M(M)$ can be used to evaluate the integrals in (49).

- Finally, inequality (49) follows from the following strings of equalities/inequalities and (50):

$$\int_0^1 p_m(t) t^{(n-2)/2} dt = \int_0^1 p_m(t) t^{N+\beta} dt = \sum_{j=1}^{\ell} \lambda_M(j) p_m(z_M(j)) z_M(j)^N$$

and

$$\begin{aligned} \int_0^1 p_m(t) t^{k+(n-2)/2} dt &= \int_0^1 p_m(t) t^{k+N+\beta} dt = \sum_{j=1}^{\ell} \lambda_M(j) p_m(z_j) z_M(j)^{k+N} \\ &\leq z_M(\ell)^k \left\{ \sum_{j=1}^{\ell} \lambda_M(j) p_m(z_j) z_M(j)^N \right\}. \end{aligned}$$

The properties of the sequence of polynomials just constructed together with the methods of Subsection 3.4 allow us to conclude the following:

Proposition 10 Suppose $P_m(x)$, $m = 1, 2, \dots$, is the family of positive radial polynomials defined by $P_m(x) = p_m(|x|^2)$ where p_m is the univariate polynomial given by (51). Define

$$\tilde{P}_m^k(x) = \sum_{j=0}^{k-1} (-1)^{k-j-1} \frac{k!}{(k-j)!j!} (k-j)^{-n} P_m(x/(k-j)).$$

Then \tilde{P}_m^k is a polynomial of degree no greater than $2m$ and for any f in $L^p(B_{1/2}^n)$, $1 \leq p \leq \infty$,

$$\|\tilde{P}_m^k * f - f\|_{L^p(B_{1/2}^n)} \leq c \omega_p^k(f; 1/m)$$

where c is a constant independent of f and m . In other words, for fixed k the sequence of polynomials $\tilde{P}_m^k * f$, $m = 1, 2, \dots$, converges to f in $L^p(B_{1/2}^n)$ at a rate which is $O(\omega_p^k(f; 1/m))$ as $m \rightarrow \infty$.

4 Incomplete data

In this section we consider summability and approximate reconstruction when the data f_u are known only for u 's in a proper subset \mathcal{A} of S^{n-1} . Applying the same philosophy used in the previous section immediately leads to the question concerning what summability kernels K can be expressed as sums of ridge functions in directions u which lie in \mathcal{A} . It is to be expected that the class of kernels which enjoy such a representation is somewhat more restricted than in the case of full data.

In the case when \mathcal{A} is a finite set the results of [37] show that the only radial kernels which can be expressed in terms of the corresponding ridge functions are polynomials. To be more specific, let $\mathcal{A} = \{u_1, \dots, u_N\}$ and suppose K is a radial function such that

$$K(x) = \sum_{i=1}^N h_i(\langle x, u_i \rangle)$$

for appropriate univariate functions h_1, \dots, h_N . Then K must be a polynomial of degree no greater than a certain bound determined by \mathcal{A} . Furthermore, this will be significant in our considerations below, the h_i 's can be chosen to be scalar multiples of one univariate function h , namely, $h_i(t) = w_i h(t)$, $i = 1, \dots, N$.

Motivated by these and related considerations we restrict our attention to radial polynomial kernels in what follows.

4.1 Quadrature on the sphere and ridge function representations

Our summability formulas are based on the following:

Proposition 11 *Suppose \mathcal{A} is a measurable subset of S^{n-1} and μ is a bounded Borel measure on \mathcal{A} . The following are equivalent:*

(i). *There is a constant c so that the formula*

$$(52) \quad \int_{\mathcal{A}} p(u) d\mu(u) = c \int_{S^{n-1}} p(u) d\sigma(u)$$

holds for all even polynomials of degree $\leq 2m$.

(ii). *Formula (52) holds for all homogeneous polynomials of degree $= 2m$.*

(iii). *There is a constant c so that the formula*

$$(53) \quad \int_{\mathcal{A}} \langle x, u \rangle^{2k} d\mu(u) = c \int_{S^{n-1}} \langle x, u \rangle^{2k} d\sigma(u)$$

holds for all x when $k = m$.

(iv). Formula (53) holds for all x and all integers k which satisfy $0 \leq k \leq m$.
In other words

$$(54) \quad \int_{\mathcal{A}} p(\langle x, u \rangle) d\mu(u) = c \int_{S^{n-1}} p(\langle x, u \rangle) d\sigma(u)$$

for every even univariate polynomial p of degree $\leq 2m$.

(v). There is a nonzero constant c such that

$$\int_{\mathcal{A}} d\mu(u) = c\sigma_n$$

and

$$\int_{\mathcal{A}} Y(u) d\mu(u) = 0$$

for every spherical harmonic Y of degree $2k$, $k = 1, \dots, m$.

(vi). For every radial polynomial P of degree $\leq 2m$ there is a univariate polynomial p with the same degree as P such that

$$(55) \quad P(x) = \int_{\mathcal{A}} p(\langle x, u \rangle) d\mu(u)$$

holds for all x . Furthermore, the formula for p in terms of P may be derived from

$$(56) \quad P(x) = \sigma_{n-1} \int_{-1}^1 p(|x|t) (1-t^2)^{(n-3)/2} dt.$$

The proof of this proposition is contained in subsection 4.1.2. Here we merely give a related definition which will be useful later.

Motivated by the above proposition we say that a subset \mathcal{A} of S^{n-1} is in $\mathcal{Q}(m)$ if there is a bounded measure μ on \mathcal{A} , which depends on \mathcal{A} of course, such that (52) holds for all polynomials homogeneous of degree $2m$, $m = 0, 1, 2, \dots$. In other words, $\mathcal{Q}(m)$ is the collection of all measurable subsets \mathcal{A} of S^{n-1} which admit a measure μ for which (ii) of Proposition 11 holds. For convenience we set $\mathcal{Q}(\infty) = \bigcap_{m=0}^{\infty} \mathcal{Q}(m)$.

The significance of $\mathcal{Q}(m)$ in our application is the fact that property (vi) of Proposition 11 holds for each of its members.

We bring to the reader's attention the following list of transparent observations concerning $\mathcal{Q}(m)$:

- S^{n-1} is contained in $\mathcal{Q}(m)$ for each m , $m = 0, 1, 2, \dots$; namely, $S^{n-1} \in \mathcal{Q}(\infty)$.
- $\mathcal{Q}(0)$ contains every non-empty subset of S^{n-1} .
- $\mathcal{Q}(m+1) \subset \mathcal{Q}(m)$ for all m .
- If $\mathcal{A} \in \mathcal{Q}(m)$ and $\mathcal{B} \supset \mathcal{A}$ then $\mathcal{B} \in \mathcal{Q}(m)$.
- If \mathcal{A} is in $\mathcal{Q}(m)$ then the measure μ associated with \mathcal{A} whose existence is guaranteed by the definition need not be unique.

4.1.1 Examples

Suppose \mathcal{A} is an open subset of S^{n-1} . It is not very difficult to see that $\mathcal{A} \in \mathcal{Q}(\infty)$.

Indeed, to see this, let $\mathcal{P}(m)$ be the space of polynomials homogeneous of degree $2m$ and let $\langle p, q \rangle_{\mathcal{A}}$ be the bilinear form defined on $\mathcal{P}(m)$ via the formula

$$\langle p, q \rangle_{\mathcal{A}} = \int_{\mathcal{A}} p(u) \overline{q(u)} d\sigma(u).$$

Clearly $\langle p, q \rangle_{\mathcal{A}}$ is a positive definite bilinear form and thus, equipped with this form as an inner product, $\mathcal{P}(m)$ is a Hilbert space. Now consider the linear form

$$\int_{S^{n-1}} p(u) d\sigma(u)$$

which is well defined on $\mathcal{P}(m)$. Elementary Hilbert space theory now implies that there is a homogeneous polynomial of degree $2m$, call it q_m , such that

$$\langle p, q_m \rangle_{\mathcal{A}} = \int_{S^{n-1}} p(u) d\sigma(u).$$

We summarize these observations as follows:

Proposition 12 *If \mathcal{A} is an open subset of S^{n-1} then $\mathcal{A} \in \mathcal{Q}(m)$ for every m , $m = 0, 1, 2, \dots$; in other words $\mathcal{A} \in \mathcal{Q}(\infty)$. Furthermore the measure μ which does the job for a given m can be taken to be $d\mu(u) = q_m(u) d\sigma(u)$ where q_m is a homogeneous polynomial of degree $2m$.*

In the case when \mathcal{F} is a finite subset of S^{n-1} (v) of Proposition 11 can be useful in determining whether \mathcal{F} is in $\mathcal{Q}(m)$.

To wit, let $\mathcal{F} = \{u_1, \dots, u_M\}$ then any measure μ on \mathcal{F} is determined by weights, or scalars, w_1, \dots, w_M and

$$\int_{\mathcal{F}} p(u) d\mu(u) = \sum_{j=1}^M p(u_j) w_j.$$

Now, if $Y_{k\ell}$, $k = 1, \dots, N(n, \ell)$, is an orthonormal basis for the space of spherical harmonics on S^{n-1} which are homogeneous of degree ℓ then by virtue of (v) of Proposition 11 we may conclude that \mathcal{F} is $\mathcal{Q}(m)$ if and only if the system of equations

$$(57) \quad \sum_{j=1}^M w_j = 1 \text{ and } \sum_{j=1}^M Y_{k\ell}(u_j) w_j = 0$$

where

$$k = 1, \dots, N(n, \ell) \text{ and } \ell = 2, 4, \dots, 2m$$

has a solution w_1, \dots, w_M .

We remind the reader that this system has

$$L(n, 2m) = \frac{(2m + n - 1)!}{(2m)!(n - 1)!}$$

equations. However the number of elements M in \mathcal{F} need not necessarily be $\geq L(n, 2m)$ in order for a solution w_1, \dots, w_M of (57) to exist. For example, in the case $n = 2$, $L(2, 2m) = 2m + 1$; but if $u_j = (\cos \theta_j, \sin \theta_j)$, $j = 1, \dots, M$, are uniformly spaced, e.g. $\theta_j = (j - 1)\pi/M$, then it suffices to have $M \geq m + 1$ to guarantee the existence of a solution for (57). This, of course, is reminiscent of classical Gaussian quadrature.

We also remark that in the case $n = 2$ any collection $\mathcal{F} = \{u_1, \dots, u_M\}$ of distinct directions is in $Q(m)$ whenever $M \geq 2m + 1$. (By distinct directions here we mean $u_i \neq \pm u_j$ whenever $i \neq j$.) In the general case it is also true that most finite subsets consisting of $\geq L(n, m)$ directions are in $Q(m)$. However, determining them is not so simple; one needs to verify (57) or some other equivalent system.

To conclude this paragraph we summarize the main result concerning finite subsets of S^{n-1} as follows:

Proposition 13 *Suppose $\mathcal{F} = \{u_1, \dots, u_M\}$ is a finite subset of S^{n-1} . Then \mathcal{F} is in $Q(m)$ if and only if the system of equations (57) has a solution w_1, \dots, w_M .*

4.1.2 Proof of Proposition 11

The proposition is an easy consequence of the following strings of implications: $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$ and $(iv) \Rightarrow (vi) \Rightarrow (iii)$.

To wit, $(i) \Rightarrow (ii) \Rightarrow (iii)$ is clear since each condition is apparently weaker than its predecessor. Applying the Laplacian in x to (53) in the case $k = m$ results in (53) in the case $k = m - 2$ and successive applications of the Laplacian show that $(iii) \Rightarrow (iv)$.

To see that $(iv) \Rightarrow (v)$, recall that any spherical harmonic of degree ℓ can be expressed as

$$Y_\ell(u) = c(n, \ell) \int_{S^{n-1}} Y_\ell(v) C_\ell^{(n-2)/2}(\langle v, u \rangle) d\sigma(v)$$

where C_ℓ^λ is the Gegenbauer polynomial of degree ℓ with respect to the weight $(1 + t^2)^{\lambda-1/2}$ and $c(n, \ell)$ is a constant which depends only on n and ℓ . Hence we may write

$$(58) \quad \int_{\mathcal{A}} Y_\ell(u) d\mu(u) = c(n, \ell) \int_{S^{n-1}} Y_\ell(v) \left(\int_{\mathcal{A}} C_\ell^{(n-2)/2}(\langle v, u \rangle) d\mu(u) \right) d\sigma(v).$$

Now, if $\ell = 2k$, $k = 1, \dots, m$, by virtue of hypothesis (iv) the inner integral on the right hand side of (58) can be replaced with

$$\int_{S^{n-1}} C_{2k}^{(n-2)/2}(\langle v, u \rangle) d\sigma(u),$$

which is equal to 0 if $k = 1, \dots, m$, and σ_n if $k = 0$. This, of course, implies the desired result.

To see that (v) \Rightarrow (i), recall that any homogeneous polynomial p of degree $2k$ can be expressed as

$$p(x) = |x|^{2k} \left(a_0 + \sum_{j=1}^k Y_{2j}(x/|x|) \right)$$

where a_0 is a constant and Y_{2j} is a spherical harmonic of degree $2j$. Hence

$$(59) \quad \int_{S^{n-1}} p(u) d\sigma(u) = a_0 \sigma_n$$

and, by virtue of (v), there is a constant c such that

$$(60) \quad \int_{\mathcal{A}} p(u) d\mu(u) = a_0 c \sigma_n.$$

A comparison of (59) with (60) and a linearity argument imply the desired result.

Finally, to see the second string of implications, observe that (iv) implies that there are positive constants a_0, \dots, a_m such that

$$\int_{\mathcal{A}} \langle x, u \rangle^{2k} d\mu(u) = a_k |x|^{2k}$$

for $k = 0, \dots, m$. Hence if P is any radial polynomial of degree $\leq 2m$, say $P(x) = \sum_{j=0}^m b_j |x|^{2j}$, then (55) must hold with $p(t) = \sum_{j=0}^m (b_j/a_j) t^{2j}$. Thus (iv) \Rightarrow (vi).

Now, if (vi) holds, equating the highest degree terms in (55) immediately implies (iii).

4.2 Summability, reconstruction, and degree of approximation

Proposition 11 together with the large variety of available polynomial summability kernels, some of which were described in Section 3.5, allow us to easily derive summability and reconstruction methods for Radon transform data of the form $\{\mathcal{R}f(u, t) : -\infty < t < \infty, u \in \mathcal{A}\}$, where \mathcal{A} is a proper subset of S^{n-1} and the phantom f is assumed to have compact support.

The following are examples of the type of reconstruction methods which are consequences of the developments outlined above:

Proposition 14 Suppose \mathcal{A} is a subset of S^{n-1} which is in $Q(\infty)$ and f is an integrable function with support in $B_{1/2}^n$. Let p_m , $m = 1, 2, \dots$, be the sequence of univariate polynomials defined by

$$p_m(t) = c_n t \left(\frac{1}{2t} \frac{d}{dt} \right)^{n-1} t^{2n-3} \int_0^1 (1-s^2)^{(n-3)/2} s^{n-1} (1-(ts)^2)^m ds.$$

Let μ_m be a measure such that

$$\int_{\mathcal{A}} P(u) d\mu_m(u) = \int_{S^{n-1}} P(u) d\sigma(u)$$

holds for all polynomials homogeneous of degree $2m$. Then

$$(61) \quad f(x) = \lim_{m \rightarrow \infty} \frac{1}{\sigma_n} \int_{\mathcal{A}} p_m * f_u(\langle x, u \rangle) d\mu_m(u)$$

for almost all x in $B_{1/2}^n$. If, in addition, f is in $L^p(B_{1/2}^n)$ for some p satisfying $1 \leq p < \infty$ then

$$\lim_{m \rightarrow \infty} \int_{B_{1/2}^n} |f(x) - \frac{1}{\sigma_n} \int_{\mathcal{A}} p_m * f_u(\langle x, u \rangle) d\mu_m(u)|^p dx = 0.$$

Furthermore, if f is uniformly continuous, (61) holds uniformly in x for $x \in B_{1/2}^n$.

Note that the summability formula (61) is a reconstruction formula for f in terms of the data $\{\mathcal{R}f(u, t) : -\infty < t < \infty, u \in \mathcal{A}\}$.

Recall that if \mathcal{A} is an open subset of S^{n-1} then the measure in the last proposition can be taken to be $d\mu_m(u) = q_m(u) d\sigma(u)$ where q_m is a polynomial homogeneous of degree $2m$. Now, if $Y_{k\ell}$, $k = 1, \dots, N(n, \ell)$, is an orthonormal basis for the space of spherical harmonics on S^{n-1} which are homogeneous of degree ℓ then the polynomial q_m may be expressed as

$$(62) \quad q_m(u) = a_0 + \sum_{\ell=2,4,\dots,2m} \sum_{k=1}^{N(n,\ell)} a_{k\ell} Y_{k\ell}(u)$$

as a function on the unit sphere where, by virtue of (iii) of Proposition 11 and the construction of q_m , the constants $a_{k\ell}$ may be calculated by solving the system of equations

$$(63) \quad a_0 \int_{\mathcal{A}} d\sigma(u) + \sum_{\ell=2,\dots,2m} \sum_{k=1}^{N(n,\ell)} a_{k\ell} \int_{\mathcal{A}} Y_{k\ell}(u) d\sigma(u) = 1$$

and

$$(64) \quad a_0 \int_{\mathcal{A}} \overline{Y_{ij}(u)} d\sigma(u) + \sum_{\ell=2,\dots,2m} \sum_{k=1}^{N(n,\ell)} a_{k\ell} \int_{\mathcal{A}} \overline{Y_{ij}(u)} Y_{k\ell}(u) d\sigma(u) = 0$$

where $i = 1, \dots, N(n, j)$ and $j = 2, 4, \dots, 2m$.

Proposition 15 *If \mathcal{A} is an open subset of S^{n-1} then the conclusions of Proposition 14 hold with $d\mu_m(u) = q_m(u)d\sigma(u)$ where q_m is a polynomial defined by (62), (63), and (64).*

Typical results concerning degree of approximation are the following:

Proposition 16 *Suppose \mathcal{A} is in $Q(m)$ and f is in $L^p(B_{1/2}^n)$ for some p , $1 \leq p \leq \infty$. Let μ be a measure such that*

$$\int_{\mathcal{A}} P(u) d\mu(u) = c \int_{S^{n-1}} P(u) d\sigma(u)$$

holds for all polynomials homogeneous of degree $2m$. Then there is a univariate polynomial Q of degree no greater than $2m$ such that

$$(65) \quad \|f(x) - \int_{\mathcal{A}} Q * f_u(\langle x, u \rangle) d\mu(u)\|_{L^p(x \in B_{1/2}^n)} \leq c\omega_p(f, 1/m)$$

where ω_p is the L^p modulus of continuity of f and c is a constant independent of \mathcal{A} , f , and m . More generally, given any positive integer k , there is a univariate polynomial Q of degree no greater than $2m$ such that

$$(66) \quad \|f(x) - \int_{\mathcal{A}} Q * f_u(\langle x, u \rangle) d\mu(u)\|_{L^p(x \in B_{1/2}^n)} \leq c\omega_p^k(f, 1/m)$$

where ω_p^k is the k -th order L^p modulus of smoothness of f and c is a constant independent of \mathcal{A} , f , and m .

Polynomials Q which do the job in the above proposition can be constructed via the methods of Subsection 3.3 combined with those of Subsection 3.5. For example, given an integer k let \hat{P}_m^k be the polynomial defined in the statement of Proposition 10. If $\hat{p}_m(|x|) = \hat{P}_m^k(x)$ then the polynomial Q defined by

$$Q(t) = c_n t \left(\frac{1}{2t} \frac{d}{dt} \right)^{n-1} t^{2n-3} \int_0^1 (1-s^2)^{(n-3)/2} s^{n-1} \hat{p}_m(ts) ds$$

does the job in (66).

Inequalities (65) and (66) can be viewed as approximate reconstruction formulas for f in terms of the data $\{\mathcal{R}f(u, t) : -\infty < t < \infty, u \in \mathcal{A}\}$. Note that the error bound holds for all f in $L^p(B_{1/2}^n)$ and depends only on f and m .

It should be mentioned that given data of the form $\{\mathcal{R}f(u, t) : -\infty < t < \infty, u \in \mathcal{A}\}$, where \mathcal{A} is in $Q(m)$ it is not possible to obtain approximate reconstruction algorithms which will give better degree of approximation than that given by (66). Thus the result in Proposition 16 is best possible

in the following sense: There are examples of \mathcal{A} in $Q(m)$ and functions f in $L^p(B_{1/2}^n)$ such that $f_u = 0$ for all u in \mathcal{A} and

$$\|f\|_{L^p/\omega_p^k}(f, 1/m) \geq c$$

where c is a constant independent of m .

The approximate reconstruction formulas described above rely on the fact that a radial polynomial $P(x)$ of degree $\leq 2m$ enjoys the ridge function representation described by (55) whenever \mathcal{A} is in $Q(m)$. It is clear that the corresponding ridge function $Q(t)$ can be obtained from (10) or one of its variants. However, for numerical work it is usually more convenient to obtain a representation of P in terms of orthogonal polynomials of the form

$$P(x) = \sum_{j=0}^m b_j P_j^{(\lambda, (n-2)/2)}(2|x|^2 - 1)$$

for some $\lambda > -1$. Then use (34) to obtain

$$Q(t) = \sum_{j=0}^m b_j C_{2j}^{\lambda+n/2}(t)$$

which can be efficiently evaluated via the three term recurrence relation. Explicit numerical examples will appear elsewhere.

5 Miscellaneous remarks

Introductory material concerning the Radon transform and its applications can be found in [8, 16, 27, 41, 47] and the pertinent references cited there. Our use of the term ridge function is adopted from [23].

The notion of representing kernels, K , in terms of ridge functions to obtain approximate reconstruction formulas has been used for quite some time under various guises, see [4, 7, 23, 47]. For example, the so-called filtered backprojection method is a concrete and numerically useful manifestation of it, see [47, 48, 49, 16, 41].

It should be clear that in order to express a convolution $K * f(x)$ in terms of the full Radon transform data of f it suffices to express K as a sum of ridge functions. For example,

$$(67) \quad K(x) = \int_{S^{n-1}} h(u, \langle x, u \rangle) d\sigma(u)$$

is such an expression; here the ridge functions vary from direction to direction. Representation (67) holds for most reasonable functions, at least in some wide sense, with h given by

$$(68) \quad h(u, \langle x, u \rangle) = (2\pi)^{-n/2} \int_0^\infty \hat{K}(ru) e^{i\langle x, u \rangle r} r^{n-1} dr$$

where \hat{K} is the n variate Fourier transform of K . Note that the pair (67) and (68) is simply a variant of the Fourier inversion formula. The restriction to uniform sums allows for an elementary development of the desired summability formulas.

The results in Sections 2.1-3.2, restricted to the case $n = 2$, together with related material and more examples were recorded and circulated in [34]; the key ideas were announced in [35]. Some of this material was published using different notation in [42, 43]; I gratefully acknowledge the referee for pointing out these references and making other constructive comments.

Radial kernels have been used in tomographic applications by other authors, most notably [20, 50]. However, the application, perspective, and basic results in [20, 50] are significantly different from ours. For example, the formula for h in terms of K is based on, in our notation,

$$\mathcal{R}(K * f)(u, t) = K_u * f_u(t)$$

and

$$f(x) = \Lambda^{n-1} \int_{S^{n-1}} \mathcal{R}f(u, \langle x, u \rangle) d\sigma(u)$$

to get

$$(69) \quad h(t) = \mathcal{R}(\Lambda^{n-1} K)(u, t) \text{ or } \Lambda^{n-1} \mathcal{R}K(u, t)$$

where Λ^{n-1} is the pseudo-differential operator whose symbol is a constant multiple of $|\xi|^{n-1}$. Note the similarity between (69) and (28). Formula (69) appears to require conditions on K which are more elaborate than those in Proposition 3 in order to make sense pointwise. Furthermore, the derivation of (69) uses an inversion formula whereas our development leads to various inversion formulas. I gratefully acknowledge Don Solmon who brought my attention to these matters and suggested the comments in this paragraph.

The uniform ridge function representation of the bivariate Poisson kernel can be found in [47]. Additional examples of uniform ridge function representations may be found in [20, 34, 50].

Observe that (5) may be regarded as a variant of

$$(\mathcal{R}^* h) * f = \mathcal{R}^*(h * \mathcal{R}f)$$

where \mathcal{R}^* is the formal adjoint of \mathcal{R} . In this formulation it has been successfully applied in other more general contexts of integral geometry, see [2, 3, 14].

Various inversion formulas are well known for full Radon transform data, see [41] and the notes and references cited there; for example, [8, 15, 16, 17, 20, 47, 50, 51].

Polynomials arise quite naturally in the study of the Radon transform, see [41] and the notes and references cited there. For example they occur in singular value decompositions [7, 24, 28, 41, 44], in optimal reconstruction:

in the L^2 sense [23], and in many other contexts [5, 8, 29, 30, 31, 32, 33, 36, 37, 39, 41]. We note that the results of [33] show that the output of any reconstruction algorithm which uses Radon transform data discretized in the angular variable and is both rotation and translation invariant in a certain mild physically meaningful sense must be the convolution of the phantom with a polynomial. The development in Section 3.3 is simply an extension of some of the results in [32].

Moduli of smoothness provide a natural and useful measure of the degree of approximation for various processes. For classical results see [12, 56]. These moduli can be used to define precise and meaningful notions of resolution for very wide classes of approximants. We will not go into the details here. However we do mention that if Ω is a sufficiently well behaved set in R^n , such as a ball or parallelepiped, then the indicator function of Ω , call it χ satisfies

$$\omega_p(\chi, \epsilon) = O(\epsilon^{1/p})$$

as ϵ goes to 0. Hence error estimates in terms of this modulus are $O(\epsilon^{1/p})$ as $\epsilon \rightarrow 0$. In the case $p = 2$ this should be compared with estimates in terms of Sobolev space type norms for such phantoms which are $O(\epsilon^\beta)$, $\beta < 1/2$; in particular, see the introductory discussion in Section IV.2 of [41].

Polynomial approximation is a classical subject, see [12, 56]. For details and various formulas concerning classical orthogonal polynomials, including Gaussian quadrature, see [11, 18, 55]. The polynomials P_m , $m = 1, 2, \dots$, described in Subsection 3.5.2 were constructed in the case $n = 2$ and $k = 1$ in [29, 30]; they are modifications of polynomials considered in [12] and used for similar purposes. Other related constructions can be found in [31].

For other approaches to the limited angle problem see [7, 41, 46] and the pertinent references cited in these works. For other types of restrictions see [9, 21, 39, 41, 45] and the pertinent references cited there. The results in Section 4.1 are extensions of some of the results in [32, 37]. The notion that a set is in $Q(m)$ appears to be related to the notion of m -resolving; see Section III.2 in [41]. For results concerning mechanical quadrature on spheres see [13, 41, 40, 54]. For properties of spherical harmonics see [11, 18, 53].

The results in Section 4.2 are an extension and improvement of some of the work recorded in [31] and announced in [32]. These estimates can be used to obtain acceptable bounds on errors due to discretization of the Radon transform in the angular variables but not in the t variable. For such bounds which take discretization in both types of variables into account but use Sobolev type norms see [41, Section IV.2] and the pertinent references cited there; note however the restrictions on the various parameters in these results. For bounds on resolution from another point of view see [22].

The degree of approximation or resolution from limited angle data depends only on the parameter m and not on any other properties of the set

A. This is not the case when conditionedness of the problem is taken into account. For estimates on the accuracy of the data required for a given resolution in the reconstruction see [38]; for results related to resolution and the size of singular values see [6, 26, 28, 41] and the pertinent references in these works.

References

- [1] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, Dover, New York, N.Y., 1965.
- [2] Bray and D. C. Solmon, The horocycle transform and harmonic analysis on the Poincare disk, preprint 1989.
- [3] Bray and D. C. Solmon, Paley-Weiner Theorems on rank one symmetric spaces of non-compact type, these proceedings.
- [4] Z. H. Cho, General views on 3-D image reconstruction and computerized transverse axial tomography, *IEEE Trans. Nucl. Sci.*, NS-21-3 (1974), 44-71.
- [5] A. M. Cormack, Representation of a function by its line integrals with some radiological applications, *J. Appl. Phys.*, 34 (1963), 2722-2727.
- [6] M. E. Davison, The ill-conditioned nature of the limited angle tomography problem, *SIAM J. Appl. Math.*, 43 (1983), 428-448.
- [7] M. E. Davison and F. A. Grunbaum, Tomographic Reconstructions with Arbitrary Directions. *Comm. Pure Appl. Math.*, 34, (1981), 77-120.
- [8] S. R. Deans, *The Radon Transform and some of its Applications*, Wiley, New York, 1983.
- [9] A. J. Devaney, The limited view problem in diffraction tomography, *Inverse Problems*, 5 (1989), 501-521.
- [10] W. F. Donoghue, *Distributions and Fourier Transforms*, Academic, New York, 1969.
- [11] A. Erdelyi, ed., *Higher Transcendental Functions*, Vol 2, Bateman Manuscript Project, McGraw Hill, New York, N.Y., 1953.
- [12] R. P. Feinerman and D. J. Newman, *Polynomial Approximation*, The Williams and Wilkens Co., Baltimore, 1974.
- [13] W. Freeden, On integral formulas of the sphere and application to numerical computation of integrals, *Computing*, 25 (1980), 131-146.

- [14] I. A. Hazou and D. C. Solmon, Approximate inversion formulas and convolution kernels for the exponential X-ray transform, *Z. angew. Math. Mech.* 66 (1986)5, 370-373.
- [15] S. Helgason, *The Radon Transform*, Birkhauser, Boston, 1980.
- [16] G. T. Herman, *Image Reconstruction from Projections. The Fundamentals of Computerized Tomography*, Academic Press, 1980.
- [17] A. Hertle, Continuity of the Radon Transform and its inverse on Euclidean Space, *Math.Z.*, 184 (1983), 165-192.
- [18] H. Hochstadt, *The Functions of Mathematical Physics*, Wiley-Interscience, New York, 1971.
- [19] S. H. Izen, Inversion of the k-plane transform by orthogonal function series expansions, *Inverse Problems*, 5 (1989), 181-202.
- [20] J. V. Leahy, K. T. Smith and D. C. Solmon, Uniqueness, Nonuniqueness and Inversion in the X-ray and Radon Problems, *Proc. International Symposium on Ill-posed Problems*, Newark, DE, October 1979.
- [21] R. M. Lewitt, R. H. T. Bates, and T. M. Peters, Image reconstruction from projections: III, Projection completion methods, *Optik*, 50 (1978), 180-205.
- [22] B. F. Logan, The Uncertainty Principle in Reconstructing Functions from Projections. *Duke Mathematical J.*, 42 (1975), 661-706.
- [23] B. F. Logan and L. A. Shepp, Optimal Reconstruction of a Function from its Projections, *Duke Math. J.*, 42 (1975), 645-659.
- [24] A. K. Louis, Picture reconstruction from projections in restricted Range, *Math. Meth. in the Appl. Sci.*, 2 (1980), 209-220.
- [25] A. K. Louis, Approximate inversion of the 3D Radon Transform. *Math. Meth. in the Appl. Sci.*, 5 (1983), 176-185.
- [26] A. K. Louis, Laguerre and computerized tomography: consistency conditions and the stability of the Radon transform, in *Polynome Orthogonaux et Application*, C. Brezinski, A. Draux, A. P. Magnus, P. Maroni, and A. Rouveaux, eds., Lecture Notes in Mathematics, Springer, 1985, 524-531.
- [27] A. K. Louis and F. Natterer, Mathematical Problems of Computed Tomography, *Proc. IEEE.*, 71 (1983), 379-389.

- [28] P. Maass, The X-ray transform: singular value decomposition and resolution, *Inverse Problems*, 3 (1987) 729-741.
- [29] W. R. Madych, Optimal reconstruction by polynomials in computerized tomography, Ames Lab Tech. Rep., 1979, 29 pp.
- [30] W. R. Madych, Degree of approximation in computerized tomography, *Approximation Theory III*, E.W. Cheney, ed., Academic, New York, 1980, 615-622.
- [31] W. R. Madych and S. A. Nelson, Polynomial based Algorithms for Computed Tomography, *SIAM J. Appl. Math.*, 43 (1983), 193-208.
- [32] W. R. Madych and S. A. Nelson, Polynomial based Algorithms for Computed Tomography, II, *SIAM J. Appl. Math.*, 44 (1984), 193-208.
- [33] W. R. Madych and S. A. Nelson, Characterization of tomographic reconstructions which commute with rigid motions, *J. Functional Analysis*, 46 (1982) 258-263.
- [34] W. R. Madych and S. A. Nelson, Stable inversion formulas in computed tomography, preprint, 1982, 39 pp.
- [35] W. R. Madych and S. A. Nelson, Approximate inversion formulas for Radon transform data, *Approximation Theory III*, C. K. Chui, L. L. Schumaker, and J. D. Ward, eds., Academic, New York, 1983, 599-604.
- [36] W. R. Madych and S. A. Nelson, Radial sums of ridge functions: a characterization, *Math. Methods Appl. Sci.*, 7 (1985), 90-100.
- [37] W. R. Madych and S. A. Nelson, Spherical quadrature and inversion of Radon transforms, *Proc. Amer. Math. Soc.*, 95 (1985), 453-457.
- [38] W. R. Madych and S. A. Nelson, Reconstruction from restricted Radon transform data: resolution and ill-conditionedness, *SIAM J. Math. Anal.*, 17, (1986), 1447-1453.
- [39] R. B. Marr, On the Reconstruction of a Function on a Circular Domain from a sampling of its Line Integrals. *J. Math. Anal. App.*, 45 (1974), 357-374.
- [40] A. D. McLaren, Optimal Numerical Integration on a Sphere, *Math. Comp.*, 17 (1963), 361-383.
- [41] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons, Stuttgart, 1986.

- [42] Y. Nievergelt, Elementary inversion of Radon's transform, *SIAM Review*, 28, (1986), 79-84.
- [43] Y. Nievergelt, Exact reconstruction filters to invert Radon transforms with finite elements, *J. Math. Anal. Appl.*, 120 (1986), 288-214.
- [44] E. T. Quinto, Singular value decompositions and inversion methods for the exterior Radon transform and a spherical transform, *J. Math. Anal. Appl.*, 95 (1983), 437-448.
- [45] E. T. Quinto, Tomographic reconstruction from incomplete data - numerical inversion of the exterior radon transform, *Inverse Problems*, 4 (1988), 867-876.
- [46] J. A. Reeds and L. A. Shepp, Limited angle reconstruction in tomography via squashing, *IEEE Trans. Med. Imaging*, MI-6 (1987), 89-97.
- [47] H. J. Scudder, Introduction to Computer Aided Tomography, *Proc. IEEE*, 66 (1978), 628-637.
- [48] L. A. Shepp and B. F. Logan, The Fourier reconstruction of a head section, *IEEE Trans. Nucl. Sci.*, NS-21-3 (1974), 24-43.
- [49] L. A. Shepp and J. B. Kruskal, Computerized tomography: the new medical X-ray technology, *Amer. Math. Monthly*, 85 (1978), 420-439.
- [50] K. T. Smith, Reconstruction formulas in Computed Tomography, in *Computed Tomography* (L.A. Shepp, ed.), *Proc. Sympos. Appl. Math.*, Vol. 27, Amer. Math. Soc., Providence, R.I., 1983 pp. 7-23.
- [51] K. T. Smith, D. C. Solmon, S. L. Wagner, Practical and Mathematical Aspects of the Problem of reconstructing a function from Radiographs, *Bull. AMS*, 83 (1977), 1227-1270.
- [52] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, N.J. 1970.
- [53] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton Univ. Press, Princeton, 1971
- [54] A. H. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall, Englewood Cliffs, N.J., 1971.
- [55] G. Szego, *Orthogonal Polynomials* Colloquium Publications, Vol. 23, Amer. Math. Soc., Providence, 1978.
- [56] A. F. Timan, *Theory of Approximation of Functions of a Real Variable*, Macmillan, New York, 1963.

Bounds on Multivariate Polynomials and Exponential Error Estimates for Multiquadric Interpolation

W. R. Madych* S. A. Nelson†

Abstract

A class of multivariate scattered data interpolation methods which includes the so-called multiquadrics is considered. Pointwise error bounds are given in terms of several parameters including a parameter d which, roughly speaking, measures the spacing of the points at which interpolation occurs. In the multiquadric case these estimates are $O(\lambda^{1/d})$ as $d \rightarrow 0$ where λ is a constant which satisfies $0 < \lambda < 1$. An essential ingredient in this development which may be of independent interest is a bound on the size of a polynomial over a cube in R^n in terms of its values on a discrete subset which is scattered in a sufficiently uniform manner.

1 Introduction

Let h be a continuous function on R^n which is conditionally positive definite of order m . Given data (x_j, f_j) , $j = 1, \dots, N$, where $X = \{x_1, \dots, x_N\}$ is a subset of points in R^n and the f_j 's are real or complex numbers, the so-called h spline interpolant of this data is the function s defined by

$$(1) \quad s(x) = p(x) + \sum_{j=1}^N c_j h(x - x_j)$$

where $p(x)$ is a polynomial in \mathcal{P}_{m-1} and the c_j 's are chosen so that

$$(2) \quad \sum_{j=1}^N c_j q(x_j) = 0$$

*University of Connecticut. Both authors were partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-86-0145

†Iowa State University.

for all polynomials q in \mathcal{P}_{m-1} and

$$(3) \quad p(x_i) + \sum_{j=1}^N c_j h(x_i - x_j) = f_i, \quad i = 1, \dots, N.$$

Here \mathcal{P}_{m-1} denotes the class of those polynomials on R^n of degree $\leq m-1$.

It is well known that the system of equations (2) and (3) has a unique solution when X is a determining set for \mathcal{P}_{m-1} and h is strictly conditionally positive definite. Thus, in this case, the interpolant $s(x)$ is well defined.

We remind the reader that X is said to be a determining set for \mathcal{P}_{m-1} if p is in \mathcal{P}_{m-1} and p vanishes on X implies that p is identically zero.

If h is the function defined by the formula

$$(4) \quad h(x) = \sqrt{1 + |x|^2}$$

where $|x|$ is the Euclidean norm of x then $m = 1$ and corresponding method of interpolation defined by (1), (2), (3) and (4) is often referred to as the multiquadric method. This and closely related methods are currently quite fashionable, see [3, 8]

In an earlier paper [6] we obtained bounds on the pointwise difference between a function f and the h spline which agrees with f on a finite subset X of R^n . These estimates involve a parameter d that measures the spacing of the points in X and are $O(d^\ell)$ as $d \rightarrow 0$ where ℓ depends on h . The results of the present paper imply that for certain h 's, which include (4), the estimates can be improved to $O(\lambda^{1/d})$ as $d \rightarrow 0$ where λ is a constant which satisfies $0 < \lambda < 1$. The conditions on f are the same as those in [6].

1.1 A bound for multivariate polynomials

A key ingredient in the development of our estimates is the following lemma which gives a bound on the size of a polynomial on a cube in R^n in terms of its values on a discrete subset which is scattered in a sufficiently uniform manner. This result may be of independent interest.

Lemma 1 *For $n = 1, 2, \dots$, define γ_n by the formulas $\gamma_1 = 2$ and, if $n > 1$, $\gamma_n = 2n(1 + \gamma_{n-1})$. Let Q be a cube in R^n that is subdivided into q^n identical subcubes. Let Y be a set of q^n points obtained by selecting a point from each of those subcubes. If $q \geq \gamma_n(k+1)$, then for all p in \mathcal{P}_k*

$$\sup_{x \in Q} |p(x)| \leq e^{2q^n} \sup_{y \in Y} |p(y)|.$$

We remark that it is not essential for the set Y to intersect every subcube of Q as hypothesized above. A variant of this lemma where Y intersects a certain percentage of these subcubes can be found in subsection 3.3.

1.2 A variational framework for interpolation

The precise statement of our estimates concerning h splines requires a certain amount of technical notation and terminology which are identical to that used in [6]. For the convenience of the reader we recall several basic notions.

The space of complex valued functions on R^n that are compactly supported and infinitely differentiable is denoted by \mathcal{D} . The Fourier transform of a function ϕ in \mathcal{D} is

$$\hat{\phi}(\xi) = \int e^{-i(x,\xi)} \phi(x) dx.$$

A continuous function h is conditionally positive definite of order m if

$$(5) \quad \int h(x) \phi * \tilde{\phi}(x) dx \geq 0$$

holds whenever $\phi = p(D)\psi$ with ψ in \mathcal{D} and $p(D)$ a linear homogeneous constant coefficient differential operator of order m . Here $\tilde{\phi}(x) = \overline{\phi(-x)}$ and $*$ denotes the convolution product

$$\phi_1 * \phi_2(t) = \int \phi_1(x) \phi_2(t-x) dx.$$

Note that (5) can be rewritten as

$$\int \int h(x-y) \phi(x) \overline{\phi(y)} dx dy \geq 0.$$

In what follows h will always denote a continuous conditionally positive definite function of order m . The Fourier transform of such distributions uniquely determines a positive Borel measure μ on $R^n \sim \{0\}$ and constants a_γ , $|\gamma| = 2m$ as follows: For all $\psi \in \mathcal{D}$

(6)

$$\int h(x) \psi(x) dx = \int \left\{ \hat{\psi}(\xi) - \hat{\chi}(\xi) \sum_{|\gamma| < 2m} D^\gamma \hat{\psi}(0) \frac{\xi^\gamma}{\gamma!} \right\} d\mu(\xi) + \sum_{|\gamma| \leq 2m} D^\gamma \hat{\psi}(0) \frac{a_\gamma}{\gamma!}$$

where for every choice of complex numbers c_α , $|\alpha| = m$,

$$(7) \quad \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha+\beta} c_\alpha \bar{c}_\beta \geq 0.$$

Here χ is a function in \mathcal{D} such that $1 - \hat{\chi}(\xi)$ has a zero of order $2m+1$ at $\xi = 0$; both of the integrals $\int_{0 < |\xi| < 1} |\xi|^{2m} d\mu(\xi)$, $\int_{|\xi| \geq 1} d\mu(\xi)$ are finite. The choice of χ affects the value of the coefficients a_γ for $|\gamma| < 2m$.

If

$$\mathcal{D}_m = \left\{ \phi \in \mathcal{D} : \int x^\alpha \phi(x) dx = 0 \text{ for all } |\alpha| < m \right\}$$

then $\mathcal{C}_{h,m}$ is the class of those continuous functions f which satisfy

$$(8) \quad \left| \int f(x)\phi(x)dx \right| \leq c(f) \left\{ \int h(x-y)\phi(x)\overline{\phi(y)}dx dy \right\}^{1/2}$$

for some constant $c(f)$ and all ϕ in \mathcal{D}_m . If $f \in \mathcal{C}_{h,m}$ let $\|f\|_h$ denote the smallest constant $c(f)$ for which (8) is true. Recall that $\|f\|_h$ is a semi-norm and $\mathcal{C}_{h,m}$ is a semi Hilbert space; in the case $m = 0$ it is a norm and a Hilbert space respectively. Elements f in $\mathcal{C}_{h,m}$ are of the form

$$f = f_1 + f_2$$

where the Fourier transform of f_1 is given by

$$\hat{f}_1(\xi) = g(\xi)d\mu(\xi)$$

with g in $L^2(d\mu)$ and f_2 is a polynomial of degree m .

Given a function f in $\mathcal{C}_{h,m}$, there is an element s of minimal $\mathcal{C}_{h,m}$ norm which is equal to f on X . If X is a determining set for \mathcal{P}_{m-1} then s is unique. We refer to such s as the h spline interpolant of f on X .

In the case when X is a finite subset of R^n as considered in beginning of this introduction the h spline s is given by (1), where $f(x_i) = f_i$, $i = 1, \dots, N$,

1.3 Exponential error estimates

Our basic theorem concerns how well s approximates f in regions Ω where X provides sufficient coverage. In other words, we are interested in bounds on the quantity

$$(9) \quad \frac{|f(x) - s(x)|}{\|f\|_h}$$

where x is in Ω ; the estimates should be in terms of parameters which measure how closely X covers Ω . For example, the parameter $d = d(\Omega, X)$ defined by

$$d(\Omega, X) = \sup_{y \in \Omega} \inf_{x \in X} |y - x|$$

is one such measure.

In [6] we showed that in many cases the quantity in (9) is $O(d^k)$ as $d \rightarrow 0$ where k is a constant whose maximum value is determined by h . In this paper we restrict our attention to h 's whose corresponding measures μ defined by (6) satisfy certain moment conditions. For example, if h is given by (4) there is a positive constant ρ such that for all integers k greater than 2

$$(10) \quad \int |\xi|^k d\mu(\xi) \leq \rho^k k!.$$

In this case we are able to obtain the exponential estimate described in the abstract.

In subsection 2.3 we consider a variant of (10) where $k!$ is replaced by k^{rk} , r an arbitrary real constant. As might be expected, this leads to somewhat different bounds on (9).

Because of the local nature of the result, we restrict our attention to the case where Ω is a cube.

Theorem 1 *Suppose h is conditionally positive definite of order m and the corresponding measure μ satisfies (10) for all k greater than $2m$. Then, given a positive number b_0 , there are positive constants δ_0 and λ , $0 < \lambda < 1$, which depend on b_0 and h for which the following is true: If $f \in C_{h,m}$ and s is the h spline that interpolates f on X then*

$$|f(x) - s(x)| \leq \lambda^{1/\delta} \|f\|_h$$

holds for all x in a cube E provided that (i) E has side b and $b \geq b_0$, (ii) $0 < \delta \leq \delta_0$, and (iii) every subcube of E of side δ contains a point of X .

Observe that every cube of side δ contains a ball of radius $\delta/2$. Thus the subcube condition is satisfied when $\delta = 2d(E, X)$. More generally, we can easily conclude the following:

Corollary 1 *Suppose h satisfies the hypotheses of the Theorem, Ω is a set which can be expressed as the union of rotations and translations of a fixed cube, and X is a subset of R^n . Then there are positive constants d_0 and λ , $0 < \lambda < 1$, which depend on b_0 and h for which the following is true: If $d \leq d_0$, $f \in C_{h,m}$ and s is the h spline that interpolates f on X then*

$$|f(x) - s(x)| \leq \lambda^{1/d} \|f\|_h$$

holds for all x in Ω where $d = d(\Omega, X)$.

Note that any ball in R^n satisfies the hypothesis on Ω in the above corollary. Indeed, any set Ω with sufficiently smooth boundary satisfies this hypothesis.

2 Details for Theorem 1, examples, and generalizations

As alluded to in the introduction, Lemma 1 is an important ingredient in the proof of this theorem. The following lemma, which is a transparent consequence of Lemma 1 and routine arguments involving linear functionals, is in convenient form for applying this ingredient.

Lemma 2 *Let Q , Y , and γ_n be as in Lemma 1. Then, given a point x in Q , there is a measure σ supported on Y such that*

$$\int p(y) d\sigma(y) = p(x)$$

for all p in \mathcal{P}_k , and

$$\int d|\sigma|(y) \leq e^{2n\gamma_n(k+1)}.$$

2.1 Proof of Theorem 1

First, let ρ , γ_n , and b_0 be the constants appearing in inequality (9), Lemma 1, and Theorem 1 respectively. Let

$$B = 2\rho\sqrt{n}e^{2n\gamma_n} \text{ and } C = \max\left\{B, \frac{2}{3b_0}\right\}.$$

Let

$$\delta_0 = \frac{1}{3C\gamma_n(m+1)},$$

where m is the order of conditional positive definiteness of h . We will show that δ_0 as defined above can be used for the constant in the statement of Theorem 1.

For now, let x be any point of the cube E and recall that Theorem 4.2 of [6] implies that

$$(11) \quad |f(x) - s(x)| \leq c_k \|f\|_h \int |y - x|^k d\sigma(y)$$

whenever $k \geq m$ where σ is any compactly supported measure such that

$$(12) \quad \int p(y) d\sigma(y) = p(x)$$

for all polynomials p in \mathcal{P}_{k-1} . Here

$$c_k = \left\{ \int \frac{|\xi|^{2k}}{(k!)^2} d\mu \right\}^{1/2}$$

whenever $k > m$ and by virtue of (9)

$$(13) \quad c_k \leq (2\rho)^k.$$

To obtain the desired bound on $|f(x) - s(x)|$ it suffices to find a suitable bound for

$$I = c_k \int |y - x|^k d\sigma(y).$$

This is done by choosing the measure σ appropriately. We proceed as follows:

Let δ be a parameter as in the statement of the Theorem. Since $\delta \leq \delta_0$ we may choose an integer k so that

$$(14) \quad 1 \leq 3C\gamma_n k \delta \leq 2.$$

Note that such a k is $\geq m+1$ and $\gamma_n k \delta \leq b_0$. Let Q be any cube which contains x , has side $\gamma_n k \delta$, and is contained in E . Subdivide Q into $(\gamma_n k)^n$ congruent subcubes of side δ . Since each of these subcubes must contain a point of X , select a point of X from each such subcube and call the resulting discrete set Y . By virtue of Lemma 1 we may conclude that there is a measure σ supported on Y which satisfies (12) and enjoys the estimate

$$(15) \quad \int d|\sigma|(y) \leq e^{2n\gamma_n k}.$$

We use this measure in (11) to obtain an estimate on I .

Using (13), (15), and the fact that support of σ is contained in Q whose diameter is $\sqrt{n}\gamma_n k \delta$ we may write

$$(16) \quad I \leq (2\rho)^k (\sqrt{n}\gamma_n k \delta)^k e^{2n\gamma_n k} \leq (C\gamma_n k \delta)^k.$$

Since

$$(C\gamma_n k \delta) \leq \frac{2}{3} \text{ and } k \geq \frac{1}{3C\gamma_n \delta}$$

inequality (16) implies that

$$I \leq \left((2/3)^{1/(3C\gamma_n)} \right)^{1/\delta}.$$

Hence we may conclude that

$$|f(x) - s(x)| \leq \lambda^{1/\delta} \|f\|_h$$

where

$$\lambda = (2/3)^{1/(3C\gamma_n)}.$$

2.2 Examples

A well known class of examples of conditionally positive definite h 's is given by

$$h(x) = \frac{\Gamma(a/2)}{(1 + |x|^2)^{a/2}}$$

where a is a fixed real number $\neq 0, -2, -4, \dots$ and Γ is the classical gamma function. The corresponding measure μ is given by

$$d\mu(\xi) = c_a |\xi|^{(a-n)/2} K_{(n-a)/2}(|\xi|)$$

where c_a is a positive constant and K_ν is a modified Bessel function of the second kind; see [6] for more details and the cases $a = 0, -2, -4, \dots$. Because of the exponential decay of $K_\nu(t)$ as $t \rightarrow \infty$ the moments of μ grow like $\rho^k k!$ and hence μ satisfies (9) whenever k is sufficiently large.

The important example of the Gaussian

$$h(x) = e^{-|x|^2}$$

has corresponding measure

$$d\mu(\xi) = (2\pi)^{n/2} e^{-|\xi|^2/4} d\xi$$

of course. The moments of μ grow like $\rho^k \sqrt{k!}$. Although Theorem 1 provides a bound on the error, in this case one expects better estimates because of the significantly slower than hypothesized growth of these moments.

More generally, consider the case when the measure μ is given by

$$d\mu(\xi) = e^{-|\xi|^a} d\xi$$

where a is a positive constant. Here, of course,

$$h(x) = \int e^{i\langle x, \xi \rangle} e^{-|\xi|^a} d\xi.$$

The moments of μ grow like $\rho^k k^{rk}$ where $r = 1/a$. The case $a = 2$ is essentially the Gaussian which together with the rest of the cases $a \geq 1$ is covered by Theorem 1. On the other hand if $0 < a < 1$ the bound on the rate of growth of the moments hypothesized in the statement of Theorem 1 fails to hold.

The theorems in Subsection 2.3 provide answers to the questions raised above.

2.3 Generalizations

As mentioned in the introduction, different bounds on the rate of growth of the moments of the measure μ result in different estimates on the difference between f and its h spline interpolant s off the interpolated set. Here we consider the case

$$(17) \quad \int |\xi|^k d\mu(\xi) \leq \rho^k k^{rk}$$

for $k > 2m$, where r is a real constant and ρ is a positive constant.

Note that in view of Stirling's formula there are positive constants ρ_1 and ρ_2 so that

$$(18) \quad \rho_1^k k^k \leq k! \leq \rho_2^k k^k.$$

Thus the case $r = 1$ was treated in Theorem 1. Also observe that Theorem 1 provides an estimate in the case $r < 1$. However it is possible to get a more sensitive estimate in this case without much more work; this is shown in Theorem 3 and its proof. We first consider the case $r \geq 1$.

Theorem 2 Suppose h is conditionally positive definite of order m and the corresponding measure μ satisfies (17) with $r \geq 1$ for all k greater than $2m$. Then, given a positive number b_0 , there are positive constants δ_0 and λ , $0 < \lambda < 1$, which depend on h , r , and b_0 and for which the following is true: If $f \in C_{h,m}$ and s is the h spline that interpolates f on X then

$$|f(x) - s(x)| \leq \lambda^{\delta^{-1/r}} \|f\|_h$$

holds for all x in a cube E provided that (i) E has side b and $b \geq b_0$, (ii) $0 < \delta \leq \delta_0$, and (iii) every subcube of E of side δ contains a point of X .

Proof In view of (17) and (18) there is a constant ρ_0 such that

$$\int |\xi|^k d\mu(\xi) \leq \rho_0^k k^{(r-1)k}.$$

Let γ_n and b_0 be the constants appearing in the statements of Lemma 1 and Theorem 2 respectively. Let

$$B = 2\rho_0 \sqrt{n} e^{2n\gamma_n} \text{ and } C = \max \left\{ B, \frac{2}{3b_0} \right\}.$$

and let

$$\delta_0 = \frac{1}{3^r C \gamma_n (m+1)^r},$$

where m is the order of conditional positive definiteness of h . Let δ be a parameter as in the statement of the theorem. Since $\delta \leq \delta_0$ so that $3(C\gamma_n\delta)^{1/r}$ is less than 1 and we may choose an integer k so that

$$1 \leq 3(C\gamma_n\delta)^{1/r} k \leq 2.$$

Note that such a k is $\geq m+1$ and $\gamma_n k \delta \leq b_0$.

Proceeding as in the proof of Theorem 1 we get

$$(19) \quad |f(x) - s(x)| \leq I \|f\|_h$$

where

$$I \leq \rho_0^k k^{(r-1)k} (\sqrt{n} \gamma_n k \delta)^k e^{2n\gamma_n k} \leq ((C\gamma_n\delta)^{1/r} k)^{rk}.$$

Since

$$(C\gamma_n\delta)^{1/r} k \leq \frac{2}{3} \text{ and } k \geq \frac{1}{3(C\gamma_n\delta)^{1/r}}$$

we may conclude that

$$I \leq \left((2/3)^{1/(3(C\gamma_n)^{1/r})} \right)^{\delta^{-1/r}}.$$

In view of (19) the theorem now follows with

$$\lambda = (2/3)^{1/(3(C\gamma_n)^{1/r})}.$$

■

Theorem 3 Suppose h is conditionally positive definite of order m and the corresponding measure μ satisfies (17) with $r < 1$ for all k greater than $2m$. Then, given a positive number b_0 , there are positive constants δ_0 , c , and C , which depend on h , r , and b_0 and for which the following is true: If $f \in C_h$ and s is the h spline that interpolates f on X then

$$|f(x) - s(x)| \leq (C\delta)^{c/\delta} \|f\|_h$$

holds for all x in a cube E provided that (i) E has side b and $b \geq b_0$, (ii) $0 < \delta \leq \delta_0$, and (iii) every subcube of E of side δ contains a point of X .

Proof Let

$$\delta_0 = \min \left\{ \frac{1}{(Bb_0^r)^{1/(1-r)}\gamma_n}, \frac{b_0}{2\gamma_n(m+1)} \right\}$$

where γ_n is the constant defined in Lemma 1, and

$$B = \rho_0 \sqrt{n} e^{2n\gamma_n}$$

with ρ_0 as in the proof of Theorem 2. Then if $\delta \leq \delta_0$ there is an integer k so that

$$\frac{b_0}{2} \leq \gamma_n \delta k \leq b_0.$$

Arguing as in the proof of Theorem 2 we can conclude that

$$(20) \quad |f(x) - s(x)| \leq I \|f\|_h$$

where $I \leq (B\gamma_n \delta k^r)^k$. Since $k \leq b_0/(\gamma_n \delta)$ we may write

$$I \leq \left(B\gamma_n \delta \left(\frac{b_0}{\gamma_n \delta} \right)^r \right)^k$$

and since $B\gamma_n^{1-r} b_0^r \delta^{1-r} \leq 1$ and $k \geq b_0/(2\gamma_n \delta)$ it follows that

$$I \leq (B\gamma_n^{1-r} b_0^r \delta^{1-r})^{b_0/(2\gamma_n \delta)}.$$

The last inequality together with (20) imply the desired result with

$$C = (Bb_0^r)^{1/(1-r)}\gamma_n \text{ and } c = \frac{(1-r)b_0}{2\gamma_n}.$$

■

3 Details for Lemma 1

We begin by noting that it suffices to prove the Lemma in the case $Q = [0, 1]^n$. To see this, let Q be any cube in R^n and let ϕ be an affine transformation mapping $[0, 1]^n$ onto Q . Then polynomials p on Q are related to polynomials f on $[0, 1]^n$ via the correspondence

$$f(x) = p(\phi(x))$$

and the corresponding subdivisions and discrete subsets Y are related analogously. It is clear that an estimate like that given by Lemma 1 on the size of f on $[0, 1]^n$ implies the corresponding estimate on the size of p on Q . Hence in what follows we will always take the cube Q to be $[0, 1]^n$.

Our proof of Lemma 1 involves induction on the dimension n . While Lemma 1 and its proof are elementary and well known in the case $n = 1$, in the first subsection we formulate it in a manner convenient for the necessary induction argument. Since the general case involves certain unpleasant combinatoric and geometric complications, for the sake of clarity we spell out the argument in the case $n = 2$ in the second subsection. The general case is considered in the third subsection.

3.1 The case $n=1$

Proposition 1 *Let $T = \{t_0, \dots, t_k\}$ be a subset of the unit interval $[0, 1]$ and assume $t_{i-1} + 1/q \leq t_i$, for $i = 1, \dots, k$. Then for all $p \in P_k$,*

$$\sup_{t \in [0, 1]} |p(t)| \leq \frac{(2q)^k}{k!} \sup_{t \in T} |p(t)|.$$

Proof Recall $p = \sum_{i=0}^k p(t_i) L_i$ where

$$L_i(t) = \prod_{j=0, j \neq i}^k \frac{t - t_j}{t_i - t_j}.$$

The assumption $1/q \leq t_i - t_{i-1}$ implies $|t_i - t_j| \leq q/|i - j|$. Also, $|t - t_j| \leq 1$ for all $t \in [0, 1]$. Hence, for such t , $|L_i(t)| \leq q^k/[i!(k-i)!]$ and

$$\sum_{i=0}^k |L_i(t)| \leq q^k \sum_{i=1}^k \frac{1}{i!(k-i)!} = \frac{(2q)^k}{k!}$$

which gives the desired inequality. ■

Since

$$\frac{(2q)^k}{k!} \leq e^{2q}$$

this is a simple variant of Lemma 1 in the case $n = 1$.

3.2 The case $n=2$

Proposition 2 Suppose the square $Q = [0, 1]^2$ is divided into q^2 identical subsquares and X is a set that intersects each subsquare. If $q \geq 12(k+1)$, then for all $p \in \mathcal{P}_k$

$$(21) \quad \sup_{x \in Q} |p(x)| \leq e^{48(k+1)} \sup_{x \in X} |p(x)|.$$

Proof Let Q_i , $i \in I$ denote the q^2 subsquares of Q . If $q > 12(k+1) \equiv q'$ then Q contains squares Q' that consist of exactly $(q')^2$ subsquares Q_i . Since (21) holds for Q if it holds for all such Q' , we assume that $q = 12(k+1)$.

Instead of (21) we will show that if $h \in \mathcal{P}_k$ and $|h(x)| < 1$ for all $x \in X$ then

$$(22) \quad \sup_Q |h| \leq e^{48(k+1)}.$$

That this implies (21) can be seen by considering $h = p/(\epsilon + \sup_X |p|)$, $\epsilon > 0$.

Let $m_i = \min_{\partial Q_i} |h|$ where ∂Q_i denotes the boundary of Q_i . Let N_0 be the number of points in $I_0 = \{i \in I : m_i < 1\}$. We assert that

$$(23) \quad N_0 \geq (12(k+1))^2 - (2k-1)^2.$$

To see this, take $b = (b_1, b_2)$, let $g_b(x) = |h(x)|^2 + (b_1 x_1 + b_2 x_2)$, and note that for every $i \in I \setminus I_0$,

$$\min_{Q_i} g_0 < 1 \leq (m_i)^2 = \min_{\partial Q_i} g_0.$$

Thus we can choose $\epsilon > 0$ so that if $|b| < \epsilon$ then for every $i \in I \setminus I_0$

$$\min_{Q_i} g_b < \min_{\partial Q_i} g_b.$$

When this occurs, g_b has a critical point in the interior of Q_i . Such a b can be chosen so that all the critical points of g_b are nondegenerate, for example see Lemma 6.2 on page 40 of [7]. Now $g_b \in \mathcal{P}_{2k}$, so by virtue of Proposition 4 in Subsection 3.4 it can have at most $(2k-1)^2$ nondegenerate critical points. Thus $I \setminus I_0$ has at most $(2k-1)^2$ points and (23) follows.

For each $i \in I_0$ select a point $y_i \in \partial Q_i$ so that $|h(y_i)| < 1$ and so that y_i is not one of the four corners of Q_i . Partition I_0 into four subsets I_1, \dots, I_4 according to whether y_i lies on the top, bottom, left or right edge of Q_i . Let N_1 be the number of points in I_1 and assume without loss of generality that $N_1 \geq N_0/4$.

For each $j = 1, \dots, q$ let $I(j)$ be the set of i 's for which Q_i lies in the horizontal strip

$$\{(t, s) : 0 \leq t \leq 1, (j-1) \leq qs \leq j\}.$$

Let $N(j)$ be the number of points in $I_1 \cap I(j)$ and let N be the number of points in $J = \{j : N(j) \geq 2(k+1)\}$.

Note that $N_1 = \sum_{j=1}^q N(j) \leq Nq + (q-N)(2k+1)$. Using $q = 12(k+1)$, this gives

$$N_1 \leq N(10k+11) + 12(k+1)(2k+1)$$

If $N \leq k$ then $N_0 \leq 4N_1 \leq 4[(10k^2 + 11k) + 12(k+1)(2k+1)]$ which is easily seen to violate (23). We conclude that $N \geq k+1$.

Let $p_j(t) = h(t, j/q)$. In $N(j)$ of the intervals

$$\frac{r-1}{q} \leq t \leq \frac{r}{q}, \quad r = 1, \dots, q$$

there is a point t with $|p_j(t)| < 1$. If $j \in J$ there are at least $2(k+1)$ such intervals. Thus we can apply Proposition 1 to p_j and see that

$$(24) \quad \max_{t \in [0,1]} |p_j(t)| \leq e^{2q} = e^{24(k+1)}$$

for every $j \in J$. Using this and the fact that J has $N \geq k+1$ points we can apply Proposition 1 again, this time to $p(s) = h(a, s)$, $a \in [0, 1]$ to arrive at (22). ■

3.3 The general case

Proposition 3 Define γ_n for $n = 1, 2, \dots$ by $\gamma_1 = 2$, $\gamma_n = 2n(1 + \gamma_{n-1})$, $n > 1$. Let X be a subset of R^n , let $r \in (0, 1]$ and let k and q be positive integers with $q \geq \gamma_n(k+1)/r$. Subdivide the unit n -cube $[0, 1]^n$ into q^n identical subcubes and let N be the number of such subcubes that intersect X . If $N \geq rq^n$ then for all $f \in \mathcal{P}_k$

$$(25) \quad \sup_{x \in [0,1]^n} |f(x)| \leq \left(\frac{(2q)^k}{k!} \right)^n \sup_{x \in X} |f(x)|$$

Proof We first deal with the case $n = 1$. In that case the subcubes are the intervals $I_i = [(i-1)/q, i/q]$, $i = 1, \dots, q$. Let $i(1) < i(2) < \dots < i(N)$ give the intervals that intersect X . For each $j = 1, \dots, N$ choose $x(j) \in I_{i(j)} \cap X$. By assumption, $N \geq rq \geq 2(k+1)$. The points

$$t_0 = x(1), t_1 = x(3), \dots, t_k = x(1+2k)$$

satisfy $t_j - t_{j-1} \geq 1/q$ so (25) follows from Proposition 1.

To complete the proof we use induction on n . The integers k and q will be held fixed during the induction. Let $n' = n-1$ and define r' by $\gamma_{n'}/r' = \gamma_n/r$. Then $q \geq \gamma_{n'}(k+1)/r'$. Subdivide the unit n' -cube $[0, 1]^{n'}$ into $q^{n'}$ identical

subcubes and let N' be the number of such subcubes that intersect $X' \subset R^{n'}$. If $N' \geq r'q^{n'}$ then, by induction, for all $g \in \mathcal{P}_k$

$$(26) \quad \sup_{[0,1]^{n'}} |g| \leq \left(\frac{(2q)^k}{k!} \right)^{n'} \sup_X |g|$$

Instead of (25) we will show that if $h \in \mathcal{P}_k$ and $|h(x)| < 1$ for all $x \in X$ then

$$(27) \quad \sup_{[0,1]^n} |h| \leq \left(\frac{(2q)^k}{k!} \right)^n.$$

That this implies (25) can be seen by considering $h = p/(\epsilon + \sup_X |p|)$, $\epsilon > 0$.

Let \mathcal{Q} denote the family of q^n subcubes of $[0,1]^n$. For each $Q \in \mathcal{Q}$ let $m_Q = \min_{\partial Q} |h|$ where ∂Q denotes the boundary of Q . Let

$$\mathcal{Q}_h = \{Q \in \mathcal{Q} : m_Q < 1\}, \mathcal{Q}_X = \{Q \in \mathcal{Q} : Q \cap X \neq \emptyset\}.$$

Note that N is the number of elements in \mathcal{Q} and let N_h be the number of elements in \mathcal{Q}_h . We assert that

$$(28) \quad N_h \geq N - (2k)^n.$$

To see this, for $b \in R^n$ consider the function g_b defined by

$$g_b(x) = |h(x)|^2 + (b_1 x_1 + \dots + b_n x_n).$$

If $Q \in \mathcal{Q}_X \setminus \mathcal{Q}_h$ then

$$\min_Q g_0 < 1 \leq \min_{\partial Q} g_0.$$

Thus we can choose $\epsilon > 0$ so that for all $Q \in \mathcal{Q}_X \setminus \mathcal{Q}_h$ and all $|b| < \epsilon$

$$\min_Q g_b < \min_{\partial Q} g_b.$$

When this holds, it is evident that g_b has a critical point in the interior of Q . Thus g_b has at least $N - N_h$ critical points. Such a b can be chosen so that the critical points of g_b are nondegenerate, see Lemma 6.2 on page 40 of [7]. Since $g_b \in \mathcal{P}_{2k}$, by virtue of Proposition 4 in Subsection 3.4 it can have at most $(2k-1)^n$ nondegenerate critical points. Thus $N - N_h \leq (2k-1)^n$ which gives (28).

For each $Q \in \mathcal{Q}_h$ a point $y(Q) \in \partial Q$ can be selected so that $|h(y(Q))| < 1$. By moving $y(Q)$ slightly, if necessary, it may also be assumed that $y(Q)$ lies on exactly one of the hyperplanes

$$M_{m,j} = \{y \in R^n : y_m = j/q\}, m = 1, \dots, n, j = 0, \dots, q.$$

Let $N_h(m, j)$ be the number of Q 's for which $y(Q) \in M_{m,j}$. Let $N_{h,m} = \sum_{j=0}^q N_h(m, j)$, and note that $N_h = \sum_{m=1}^n N_{h,m}$. Without loss of generality we assume $N_{h,n} \geq N_h/n$.

Let $Y = \{y(Q) : Q \in \mathcal{Q}_h\}$, $Y_j = Y \cap M_{n,j}$. In each hyperplane $M_{n,j}$ there are q^{n-1} $(n-1)$ -cubes that correspond to the subdivision of $[0, 1]^n$ into q^n n -cubes. Let $N(Y_j)$ be the number of $(n-1)$ -cubes in $M_{n,j}$ that intersect Y_j . Then $N(Y_j) \geq N_h(n, j)/2$ because for each $(n-1)$ -cube Q' in $M_{n,j}$ there are at most two n -cubes $Q \in \mathcal{Q}$ which contain Q' . Thus we have

$$(29) \quad 2\left(\sum_{j=0}^q N(Y_j)\right) \geq N_{h,n} \geq N_h/n.$$

If $N(Y_j) \geq r'q^{n-1}$ then from (26) we get

$$(30) \quad |h(x', j/q)| \leq \left(\frac{(2q)^k}{k!}\right)^{n-1} \sup_{Y_j} |h| < \left(\frac{(2q)^k}{k!}\right)^{n-1}$$

for all $x' \in [0, 1]^{n-1}$. Let $J = \{j : N(Y_j) \geq r'q^{n-1}\}$. We will show below that J has at least $k+1$ elements. This allows us to apply Proposition 1 to $p(t) = h(x', t)$. The result is

$$|h(x', t)| \leq \frac{(2q)^k}{k!} \max_{j \in J} |h(x', j/q)|$$

for every $t \in [0, 1]$. Because of (30), this gives (27).

Let s be the number of elements in J . It remains to show $s \geq k+1$. For all j , $N(Y_j) \leq q^{n-1}$ and for $j \notin J$, $N(Y_j) < r'q^{n-1}$. Thus

$$\sum_{j=0}^q N(Y_j) \leq sq^{n-1} + (1+q-s)r'q^{n-1}.$$

Combining this with (30), (28) and the hypothesis $N \geq rq^n$ gives

$$\frac{1}{2n}(rq^n - (2k)^n) \leq \sum_{j=0}^q N(Y_j) \leq sq^{n-1} + (1+q-s)r'q^{n-1}$$

or, after division by q^{n-1} ,

$$(31) \quad \frac{rq}{2n} - \frac{(2k)^n q}{q^n 2n} - (1+q)r' \leq s(1-r')$$

By definition of r' , $r = r'\gamma_n/\gamma_{n-1}$ with $\gamma_n = 2n(1+\gamma_{n-1})$. Hence $r/2n = r'(1+\gamma_{n-1})/\gamma_n$ or $r/2n - r' = r'/\gamma_{n-1}$. Thus (31) can be rewritten as

$$\frac{r'q}{\gamma_{n-1}} - \left(\frac{(2k)^n q}{q^n 2n} + r'\right) \leq s(1-r').$$

By assumption we have $q \geq \gamma_n(k+1)/r = \gamma_{n-1}(k+1)/r'$. Taking $M = \gamma_{n-1}/r'$ in the following lemma, we find $(1-r')(k+1) \leq s(1-r')$ which gives $s \geq k+1$. ■

Lemma 3 If $n \geq 2$, $k \geq 1$, $r' \in (0, 1]$, $Mr' \geq 2$ and $q \geq M(k+1)$ then

$$(32) \quad (1 - r')(k+1) \leq \frac{q}{M} - \frac{(2k)^n q}{q^n 2n} - r'.$$

Proof From $k \leq k+1 \leq q/M$ we have $k/q \leq 1/M \leq 1/2$ and

$$\left(\frac{2k}{q}\right)^n \frac{M(k+1)}{2n} \leq \left(\frac{2}{M}\right)^2 \frac{M(k+1)}{2n} \leq \frac{k+1}{M} \leq \frac{2k}{M} \leq kr'.$$

Multiplying this by -1 and then adding $1+k$ gives

$$1 + k - kr' \leq \left(\frac{1}{M} - \left(\frac{2k}{q}\right)^n \frac{1}{2n}\right) M(k+1).$$

Hence

$$1 + k - kr' \leq \left(\frac{1}{M} - \left(\frac{2k}{q}\right)^n \frac{1}{2n}\right) q$$

which is the same as (32). ■

3.4 Critical points of polynomials

Proposition 4 If p is a real valued polynomial on R^n of degree d then p can have at most $(d-1)^n$ nondegenerate critical points.

Proof A simple argument for the case $n = 2$ goes as follows: Let q be the greatest common factor of $\partial p / \partial x_1$ and $\partial p / \partial x_2$, and write $\partial p / \partial x_i = qp_i$, $i = 1, 2$. If q vanishes at x_0 then

$$\frac{\partial^2 p}{\partial x_i \partial x_j}(x_0) = p_j(x_0) \frac{\partial q}{\partial x_i}(x_0), \text{ so } \det \left(\frac{\partial^2 p}{\partial x_i \partial x_j}(x_0) \right) = 0.$$

Hence x_0 is a degenerate critical point. At any nondegenerate critical point x_0 we therefore have $p_1(x_0) = 0 = p_2(x_0)$. Since p_1 and p_2 have no common factor, the two variable version of Bezout's theorem, for example see [10], implies that the number of such points x_0 does not exceed $N = (\deg p_1)(\deg p_2) \leq (d-1)^2$. The lack of such a convenient form of Bezout's theorem when $n > 2$ is what makes the general case more difficult.

To obtain a proof in the general case we begin by observing that it is a corollary of its complex analogue. Indeed, there is a unique $P \in \mathcal{P}_d(C^n)$ such that $p(x) = P(x + \iota 0)$ for all $x \in R^n$. Here and in what follows $\iota = \sqrt{-1}$. From

$$\frac{\partial p}{\partial x_k}(x) = \frac{\partial P}{\partial z_k}(x + \iota 0)$$

and the corresponding formula for second order partial derivatives, it is clear that if x_0 is a nondegenerate critical point of p then $z_0 = x_0 + \iota 0$ is a nondegenerate critical point of P . Thus the general case follows from next proposition. ■

Proposition 5 *If $p \in \mathcal{P}_d(C^n)$ then p can have at most $(d-1)^n$ nondegenerate critical points.*

Proof For $j = 1, \dots, n$ let $p_j = \partial p / \partial z_j$. All critical points of p are degenerate if $p_j = 0$, so we assume $p_j \neq 0$ for all j . Let $m = \dim \mathcal{P}_d(C^n)$; we identify points $c \in C^m$

$$(33) \quad c = (c_\alpha)_{|\alpha| \leq d} = (a_\alpha + \iota b_\alpha)_{|\alpha| \leq d} = a + \iota b$$

with points $(a, b) \in R^{2m}$. For $z_0 \in C, z \in C^n$ and $c \in C^m$ let

$$f(z_0, z, c) = \sum_{|\alpha| \leq d} c_\alpha z^\alpha z_0^{d-|\alpha|}.$$

Let c_p be the point in C^m such that $p(z) = f(1, z, c_p)$ for all $z \in C^n$. Note that $p_j(z) = f_j(1, z, c_p)$ where $f_j = \partial f / \partial z_j, j = 1, \dots, n$.

Let $z(1), \dots, z(N)$ be nondegenerate critical points of p . Put $\xi^{(r)} = (1, z(r)), r = 1, \dots, N$ and observe that $\mathbf{z} = \lambda \xi^{(r)}, \lambda \in C$ is a solution of system $f_j(\mathbf{z}, c_p) = 0, j = 1, \dots, n$. By Bezout's Theorem, [9], if n homogeneous equations $f_j(\mathbf{z}) = 0$, in $n+1$ variables $\mathbf{z} = (z_0, z)$ have only a finite number of solution rays $\mathbf{z} = \lambda \xi^{(r)}, r = 1, \dots, q, \xi^{(r)} \in C^{n+1} \setminus \{0\}$, then $q \leq (d-1)^n$ where $d-1$ is the degree of $f_j, j = 1, \dots, n$. The desired conclusion, $N \leq (d-1)^n$, would follow if we knew that the system $f_j(\mathbf{z}, c_p) = 0, j = 1, \dots, n$ had only a finite number of solution rays. The latter may not be true, but it suffices to show that we can perturb c_p to obtain a point $c \in C^m$ for which the number, q_c , of solution rays of the system $f_j(\mathbf{z}, c) = 0, j = 1, \dots, n$ is finite and satisfies $q_c \geq N$.

First we show that $q_c \geq N$ is automatic if c is close enough to c_p . Consider the map T from $C^n \times C^m$ to C^n given by

$$T(z, c) = (f_1(1, z, c), \dots, f_n(1, z, c)).$$

The points $z(i)$ are nondegenerate, so the $n \times n$ matrix $\partial T / \partial z$ is nonsingular at $(z(i), c_p), i = 1, \dots, N$. By the Implicit Function Theorem there are analytic functions ζ_i on a neighborhood $B \subset C^m$ of c_p such that

$$T(\zeta_i(c), c) = 0, \zeta_i(c_p) = z(i), i = 1, \dots, N.$$

By making B smaller, if necessary, it may be assumed that $\zeta_i(c) \neq \zeta_j(c)$ for all $c \in B$ and all $i \neq j$. It is then evident that $q_c \geq N$ for all $c \in B$.

To complete the proof we will establish that for almost every point $(a, b) \in R^{2m}$, the system $f_j(\mathbf{z}, a + \iota b) = 0, j = 1, \dots, n$ has only a finite number of solution rays. For $k = 0, \dots, n$ define maps J^k from R^{2n} to $\{\mathbf{z} \in C^{n+1} : \mathbf{z}_k = 0\}$ by

$$J^k(x_1, \dots, 2n) = (x_1 + \iota x_n, \dots, x_k + \iota x_{n+k}, 1, x_{k+1} + \iota x_{n+k+1}, \dots, x_n + \iota x_{2n}).$$

Let $V(k, a, b) = \bigcap_{j=1}^n \{x \in R^{2n} : f_j(J^k(x), a + ib) = 0\}$. The maps J^k provide coordinate systems for complex projective n space. By compactness of that space, it suffices to prove that $V(k, a, b)$ consists of isolated points for every $k = 0, \dots, n$ and almost all $(a, b) \in R^{2n}$.

The proof of this uses a theorem from [7]. To prepare, define $f_{j,\alpha}(z)$, $|\alpha| \leq d$ by

$$f_{j,\alpha}(z_0, z) = (z_0)^{d-|\alpha|} \frac{\partial z^\alpha}{\partial z_j} = \frac{\partial f_j}{\partial c_\alpha}(z_0, z, c)$$

and identify $f_{j,\alpha}$ with an $n \times m$ matrix.

We assert that $f_{j,\alpha}(J^k(x))$ has rank n for every $k = 0, \dots, n$ and every $x \in R^{2n}$. For $k \neq 0$ we take

$$\alpha = \alpha(i, k) = e(i) + (d-1)e(k), \quad i = 1, \dots, n$$

where $\{e(1), \dots, e(n)\}$ is the standard basis for R^n and consider the $n \times n$ matrix $F_{j,i}(x, k) = f_{j,\alpha(i,k)}(J^k(x))$. Then $F_{k,k}(x, k) = d$, $F_{j,j}(x, k) = 1$ for $j \neq k$ and $F_{j,i}(x, k) = 0$ for $j \neq i, i \neq k$. It follows that $\det(F_{j,i}(x, k)) = d$, $k \neq 0$; the off diagonal entries of the k^{th} column of F are not needed for this. For $k = 0$, the $n \times n$ matrix $F_{j,i}(x, 0) = f_{j,e(i)}(J^0(x))$ is seen to be $\delta_{i,j}$ and our assertion is verified.

To obtain notation more like [7] we fix $k \in \{0, \dots, n\}$ and define real valued functions U_1, \dots, U_{2n} by

$$U_j(x, a, b) + iU_{j+n}(x, a, b) = f_j(J^k(x), a + ib).$$

Using the analysis of $F_{j,i}(x, k)$ above, we see that the $2n \times 2(n+m)$ matrix of partial derivatives of U_1, \dots, U_{2n} has rank $2n$. By Theorem 7.1 on page 50 of [7] we conclude that for almost all $(a, b) \in R^{2n}$, the $2n \times 2n$ matrix $\frac{\partial U_i}{\partial x_j}(x, a, b)$ is nonsingular at every point in

$$V(k, a, b) = \bigcap_{i=1}^{2n} \{x \in R^{2n} : U_i(x, a, b) = 0\}.$$

Thus for such (a, b) the points in $V(k, a, b)$ are isolated. ■

4 Miscellaneous remarks

Note that it follows from Lemma 1 that Y is a determining set for \mathcal{P}_k . We also note that in the cases where Y is regularly distributed in Q , for example if it consists of the centers of each of the subcubes, then the lemma can be derived by more traditional methods. Indeed, in the example mentioned above, it is an easy consequence of the results in [1].

The analogues of Corollary 1 for Theorems 2 and 3 are clear. It is also clear that the analogues of Lemma 1 and the Theorems hold when the cubes

are replaced by more general parallelepipeds; simply apply an appropriate affine transformation. Thus analogues of Corollary 1 hold when Ω satisfies an interior cone condition. Since our results seem to apply to most reasonable situations we refrain from exploring further generalizations.

If the measure μ satisfies (17) with $r \leq 0$ then it must have compact support. Also recall that in this case the constant C can be taken to be independent of b_0 . Since the exponent c is $(1-r)b_0/(2\gamma_n)$, if δ is such that $C\delta < 1$, letting $b_0 \rightarrow \infty$ it is clear that $|f(x) - s(x)| \rightarrow 0$. In other words, for sufficiently small δ if the intersection of X with any cube of side δ is not empty then $s(x) = f(x)$ on R^n . This means, of course, that the values of f on X uniquely determine f . The implications of this to sampling theory, such as that found in [2] or [4] for example, will be explored elsewhere.

References

- [1] L. P. Bos, Bounding the Lebesgue function for Lagrange interpolation in a simplex, *J. of Approx. Theory*, 38 (1983), 43-59.
- [2] P. L. Butzer and G. Hinsen, Two dimensional nonuniform sampling expansions - an iterative approach, I and II, *Applicable Analysis*, 32 (1989), 53-85.
- [3] N. Dyn, Interpolation and approximation by radial and related functions, in *Approximation Theory VI*, Vol. I (C. K. Chui, L. L. Schumaker, J. D. Ward, eds.), Academic Press, Boston, 1989, pp. 211-234.
- [4] H. G. Feichtinger and K. Grochenig, Reconstruction of band limited functions from irregular sampling values, preprint.
- [5] W. R. Madych and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions, *Approx. Theory and its Applications*, 4 (1988), no.4, 77- 89.
- [6] W. R. Madych and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions. II, *Math. Comp.*, 54 (1990), 211-230.
- [7] M. Morse and S. S. Cairns, *Critical Point Theory in Global Analysis and Differential Topology*, Academic Press, New York, 1969.
- [8] F. J. Narcowich and J. D. Ward, Norms of inverses and condition numbers for matrices associated with scattered data, preprint.
- [9] B. L. van der Waerden, *Modern Algebra*, vol. 2, Ungar, New York, 1964.
- [10] R. J. Walker, *Algebraic Curves*, Dover, New York, 1962.

Error estimates for interpolation by generalized splines

W. R. Madych*
Department of Mathematics, U-9
University of Connecticut
Storrs, CT 06269

Abstract

Interpolation with linear combinations of translates of a conditionally positive definite function h are considered. Such interpolants are solutions to a constrained variational problem involving the data and a quadratic form determined by h . In the case where the data is the restriction of an appropriate function f to a subset X in R^n , pointwise estimates, which are valid in a neighborhood of X , are given on the difference between f and its interpolant.

1 Introduction

Suppose h is a continuous function on R^n which is conditionally positive definite of order m . The reader not familiar with this notion should restrict his attention to the case $m = 0$ and simply take h to be the Fourier transform of a positive integrable function, e.g. $h(x) = e^{-|x|^2}$ or $h(x) = (1 + |x|^2)^{-(n+1)/2}$; the case $m = 1$ includes the examples $h(x) = |x|$ and the celebrated multiquadric $h(x) = \sqrt{1 + |x|^2}$. Such functions h need not be radial. Detailed definitions can be found in Section 2. More examples and citations can be found in Section 4.

Given data (x_j, f_j) , $j = 1, \dots, N$, where the f_j 's are scalars and the x_j 's are distinct points in R^n , consider interpolants of the form

$$(1) \quad s(x) = p(x) + \sum_{j=1}^N c_j h(x - x_j)$$

where p is a polynomial in \mathcal{P}_{m-1} , the space of polynomials of degree $\leq m-1$, and c_j 's are scalars such that

$$(2) \quad p(x_k) + \sum_{j=1}^N c_j h(x_k - x_j) = f_k, \quad k = 1, \dots, N$$

$$(3) \quad \sum_{j=1}^N c_j x_j^\alpha = 0 \quad \text{for all } \alpha, |\alpha| \leq m-1.$$

We call such interpolants *h-splines*. We also remind the reader that standard multi-index notation is used in (3). For convenient reference in what follows we denote the set of points $\{x_j\}_{j=1}^N$ by the X , namely, $X = \{x_j\}_{j=1}^N$. Note that in

*Partially supported by a grant from the Air Force Office of Scientific Research, AFOSR-89-0524

the case of $m = 0$ the polynomial p does not appear in (1); in other words, in this case $p \equiv 0$ and (3) is vacuous.

Assuming that the solution of (2)-(3) exists (it always does in the examples cited above) and the values $\{f_j\}_{j=1}^N$ are the restrictions of an appropriate function f to the set X , we are interested in estimates of the error $|f(x) - s(x)|$ for x in some open neighborhood Ω of X . Of particular interest is the asymptotic behavior of the estimate as the neighborhood Ω stays fixed and coverage by X increases.

It should be clear that to obtain bounds on the error certain restrictions on f are necessary. The restriction considered below seems to be a natural one, namely, f is taken to be such that the corresponding h -spline s is the "minimum norm" interpolant of f on X in a sense which is made more precise in Section 2. Indeed, in the estimates given below, it is not required that the set X be finite or discrete, only that s be the appropriate minimum norm interpolant of f on X .

2 Variational formulation

In what follows all integrals are taken over all of R^n , \mathcal{D} is the class of infinitely differentiable functions with compact support,

$$\mathcal{D}_m = \{\phi \in \mathcal{D} : \int x^\alpha \phi(x) dx = 0 \text{ for all } \alpha, |\alpha| < m\},$$

and the Fourier transform \hat{f} of an integrable function f is defined by

$$\hat{f}(\xi) = \int f(x) e^{-i(\xi, x)} dx.$$

All Fourier transforms are to be interpreted in the distributional sense.

A continuous function h on R^n is said to be *conditionally positive definite of order m* if and only if

$$(4) \quad \int \int h(x-y) \phi(x) \overline{\phi(y)} dx dy \geq 0$$

for all ϕ in \mathcal{D}_m . The class $CCPD_m$ is the class of all continuous positive definite functions on R^n .

If h is in $CCPD_m$ then h is a tempered distribution whose Fourier transform \hat{h} is the sum of a positive Radon measure $d\mu$ on $R^n \setminus \{0\}$ and a distribution supported at the origin. The measure is such that $\int (|\xi|/(1+|\xi|))^{2m} d\mu < \infty$ and the distribution supported at the origin is of the form $P(D)\delta$ where δ is the Dirac measure at the origin and $P(D) = \sum_\gamma a_\gamma D^\gamma$ is a partial differential operator of order $\leq 2m$ whose principle part, terms homogeneous of degree $2m$, satisfies

$$\sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha+\beta} c_\alpha \overline{c_\beta} \geq 0$$

for every choice of complex numbers $\{c_\alpha\}_{|\alpha|=m}$. In other words if ϕ is in \mathcal{D}_m then

$$(5) \quad \int \int h(x-y) \phi(x) \overline{\phi(y)} dx dy = \int |\hat{\phi}(\xi)|^2 d\mu + \|\hat{\phi}^{(m)}(0)\|_A^2$$

where

$$\|\hat{\phi}^{(m)}(0)\|_A^2 = \sum_{|\alpha|=m} \sum_{|\beta|=m} a_{\alpha+\beta} \frac{D^\alpha \hat{\phi}(0)}{\alpha!} \overline{\frac{D^\beta \hat{\phi}(0)}{\beta!}}.$$

If h is in $CCPD_m$ then the class C_h is defined as follows: f is in C_h if and only if f is a tempered distribution and there is a constant C such that

$$(6) \quad |\langle f, \phi \rangle| \leq C \left\{ \int \int h(x-y) \phi(x) \overline{\phi(y)} dx dy \right\}^{1/2}$$

for all ϕ in \mathcal{D}_m . The semi-norm $\|f\|_h$ is the infimum of all C 's which do the job in (6).

In view of (5) it is clear that (6) is equivalent to

$$|\langle f, \phi \rangle| \leq \|f\|_h \left\{ \int |\phi(\xi)|^2 d\mu + \|\phi^{(m)}(0)\|_A^2 \right\}^{1/2},$$

from which it is not difficult to conclude the following:

- The null space of the semi-norm $\|\cdot\|_h$ is \mathcal{P}_{m-1} , the class of polynomials of degree $\leq m-1$
- If f is in C_h then $f = f_1 + f_2$ where

$$(7) \quad \hat{f}_1 = g d\mu \text{ with } g \text{ in } L^2(R^n, d\mu)$$

and f_2 is a polynomial of degree $\leq m$. Of course (7) means that $\int |g(\xi)|^2 d\mu$ is finite and for any test function ϕ we have $\langle \hat{f}_1, \phi \rangle = \int \overline{\phi(\xi)} g(\xi) d\mu$.

- If $d\mu$ is absolutely continuous with respect to Lebesgue measure, namely $d\mu = w(\xi) d\xi$, and $P(D)$ has no terms of order $2m$ then

$$(8) \quad \|f\|_h^2 = \sum_{|\alpha|=m} \frac{m!}{\alpha!} \int |\xi^\alpha \hat{f}(\xi)|^2 \frac{1}{|\xi|^{2m} w(\xi)} d\xi$$

and C_h can be characterized as the class of those distributions f for which (8) is finite. Formula (8) follows from the fact that in this case $\hat{f}_1(\xi) = g(\xi)w(\xi)$ and f_2 is a polynomial of degree $\leq m-1$ so that $\xi^\alpha \hat{f}(\xi) = \xi^\alpha \hat{f}_1(\xi)$ if $|\alpha| \geq m$ and

$$\begin{aligned} \int |g(\xi)|^2 d\mu &= \int |g(\xi)w(\xi)|^2 w(\xi)^{-1} d\xi = \\ &= \sum_{|\alpha|=m} \frac{m!}{\alpha!} \int |\xi^\alpha g(\xi)w(\xi)|^2 (|\xi|^{2m} w(\xi))^{-1} d\xi. \end{aligned}$$

- If f is in C_h then f is continuous on R^n .
- If ν is a finite measure with compact support which annihilates polynomials of degree $\leq m-1$ then $\nu * h$ is in C_h . Recall that

$$\nu * h(x) = \int h(x-y) d\nu(y)$$

and ν is said to annihilate \mathcal{P}_{m-1} if and only if $\int p(x) d\nu(x) = 0$ for all p in \mathcal{P}_{m-1} .

Suppose h is in $CCPD_m$, f is in C_h , and X is a closed subset of R^n which is unisolvent for \mathcal{P}_{m-1} . (Recall that X is said to be unisolvent for \mathcal{P}_{m-1} if and only if the only polynomial in \mathcal{P}_{m-1} which vanishes on X is the constant 0.)

Then the *minimal h norm interpolant of f on X* is the unique element s in C_h which satisfies

$$(9) \quad s(x) = f(x) \text{ for all } x \text{ in } X$$

and

$$(10) \quad \|s\|_h = \min\{\|g\|_h : g \in C_h \text{ and } g(x) = f(x) \text{ for all } x \text{ in } X\}.$$

As a consequence

$$(11) \quad \|f\|_h^2 = \|f - s\|_h^2 + \|s\|_h^2$$

whenever s is a minimal h norm interpolant of f .

As mentioned in the introduction the main results considered here concern estimates of the error $|f(x) - s(x)|$ for x in a neighborhood Ω of X . Such results are essentially of a local nature and depend upon, among other things, how well points in Ω are approximated by points in X .

Suppose

$$(12) \quad \text{the inequality (4) is strict for all non-zero } \phi \text{ in } \mathcal{D}_m$$

and $X = \{x_1, \dots, x_N\}$ is a finite subset of R^n which is a uni-solvent for \mathcal{P}_{m-1} . In this case, given any collection of complex numbers $\{f_1, \dots, f_n\}$ there is an f in C_h such that $f(x_j) = f_j$, $j = 1, \dots, N$, and the minimal h norm interpolant of f on X is the h -spline defined by (1)-(3). Note that (12) is implied by a mild condition on the support of h and that all the examples mentioned above enjoy (12).

3 Error estimates

On what follows we always assume that

- h is in $CCPD_m$
- f is in C_h
- X is a closed subset of R^n
- s is the minimal h norm interpolant of f on X .

Our first estimate includes the quantity $E_k(h, \epsilon)$ which is defined by

$$E_k(h, \epsilon) = \inf_{p \in \mathcal{P}_k} \{ \sup_{|x| < \epsilon} |h(x) - p(x)| \}$$

Theorem 1 Suppose Ω is a cube of side b , $b \geq b_0 > 0$, and $k \geq m - 1$. Then there is a positive constant $\delta_0 = \delta_0(b_0, k)$ such that if every subcube of Ω of side δ where δ satisfies $0 < \delta < \delta_0$, contains a point of X then

$$(13) \quad |f(x) - s(x)| \leq C_1 \|f\|_h \{E_k(h, C_2 \delta)\}^{1/2}$$

for all x in Ω where C_1 and C_2 are positive constants which are independent of f and δ .

The proof of Theorem 1 can be found in [6]. The next result requires further hypothesis on h . More specifically, if μ is the Radon measure associated with h as in (5) then we will assume that for sufficiently large k

$$(14) \quad \int |\xi|^k d\mu(\xi) \leq \rho^k k^{\tau k}$$

where $\tau \geq 0$ and $\rho > 0$ are fixed constants.

Theorem 2 Suppose the Radon measure μ associated with h satisfies (14) for some fixed constants $r \geq 0$ and $\rho > 0$ and all integers $k \geq k_0$. Also suppose Ω is a cube of side b , $b \geq b_0 > 0$. Then there are positive constants $\delta_0 = \delta_0(b_0, h)$ and $\lambda = \lambda(b_0, h)$, $0 < \lambda < 1$, such that if every subcube of Ω of side δ where δ satisfies $0 < \delta < \delta_0$, contains a point of X then

$$(15) \quad |f(x) - s(x)| \leq \lambda^{\delta^{-a}} \|f\|_h$$

for all x in Ω where $a = \min\{1, 1/r\}$. Furthermore, if $0 \leq r < 1$, there are positive constants $c = c(b_0, h, r)$ and $C = C(b_0, h, r)$ such that

$$(16) \quad |f(x) - s(x)| \leq (C\delta)^{c/\delta} \|f\|_h$$

for all x in Ω whenever δ is sufficiently small.

The proof of this theorem in the general n variate case is technically rather complicated. For simplicity and clarity we outline here the complete argument in the univariate case which contains all the essential ingredients.

Proof To see (15) without loss of generality we may and do assume that $r \geq 1$. Let x be any point in the interval Ω , set $\phi = f - h$, let δ_0 be a positive constant which will be specified later, and suppose δ is a positive number $\leq \delta_0$. Suppose k is any positive integer such that $2k\delta \leq b$ and let Ω_x be any subinterval of Ω which contains x and has length $(2k-1)\delta$. Subdivide Ω_x into $2k-1$ subintervals of length δ ; namely,

$$\Omega_x = \cup_{l=1}^{2k-1} [a_l, a_{l+1}]$$

where $a_{l+1} - a_l = \delta$ for $l = 1, \dots, 2k-1$. Let x_j be any point in the intersection of X and the subinterval $[a_{2j-1}, a_{2j}]$, $j = 1, \dots, k$.

If p is the polynomial of degree $k-1$ which interpolates ϕ on $\{x_1, \dots, x_k\}$ then Kowalewski's exact remainder formula for polynomial interpolation reads

$$(17) \quad \phi(x) - p(x) = \frac{1}{(k-1)!} \sum_{j=1}^k \ell_j(x) \int_{x_j}^x (x_j - t)^{k-1} \phi^{(k)}(t) dt$$

where the ℓ_j 's are the Lagrange interpolating polynomials and $\phi^{(k)}$ denotes the derivative of ϕ of order k ; see [1, page 72] for details. Since $p \equiv 0$, we may estimate $\phi(x)$ by manipulating the right hand side of (17) to obtain

$$(18) \quad |\phi(x)| \leq \frac{1}{k!} \sum_{j=1}^k |\ell_j(x)| |x_j - x|^k \|\phi^{(k)}\|_\infty$$

where $\|\cdot\|_\infty$ denotes the L^∞ norm.

To estimate $\|\phi^{(k)}\|_\infty$ observe that if $k \geq m+1$ the polynomial part of ϕ is annihilated by the derivative of order k . In view of (7) for sufficiently large k we may write

$$\phi^{(k)}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi t} (i\xi)^k \hat{\psi}(\xi) d\mu(\xi)$$

where ψ is in $L^2(R^n, d\mu)$. Applying Schwartz's inequality to the last formula results in

$$(19) \quad \|\phi^{(k)}\|_\infty \leq \frac{1}{2\pi} \left\{ \int_{-\infty}^{\infty} \xi^{2k} d\mu(\xi) \right\}^{1/2} \left\{ \int_{-\infty}^{\infty} |\hat{\psi}(\xi)|^2 d\mu(\xi) \right\}^{1/2}.$$

Now

$$\left\{ \int_{-\infty}^{\infty} \xi^{2k} d\mu(\xi) \right\}^{1/2} \leq (2\rho)^k k^{rk} \text{ and } \left\{ \int_{-\infty}^{\infty} |\tilde{\psi}(\xi)|^2 d\mu(\xi) \right\}^{1/2} = \|\phi\|_h,$$

and, by virtue of (11), $\|\phi\|_h \leq \|f\|_h$. Thus (19) simplifies to

$$(20) \quad \|\phi^{(k)}\|_{\infty} \leq \frac{1}{2\pi} (2\rho)^k k^{rk} \|f\|_h.$$

To estimate the remaining terms in (18) recall that

$$\ell_j(x) = \prod_{i \neq j} \frac{x - x_i}{x_j - x_i}, \quad |x - x_j| \leq 2k\delta, \text{ and } |x_j - x_i| \geq |j - i|\delta.$$

By virtue of these inequalities we may write

$$(21) \quad \sum_{j=1}^k |\ell_j(x)| |x_j - x|^k \leq e^{4k} (2k\delta)^k.$$

Finally, substitute (20) and (21) into (19) to get

$$|\phi(x)| \leq \frac{1}{k! 2\pi} e^{4k} (2k\delta)^k (2\rho)^k k^{rk} \|f\|_h.$$

This expression can be simplified by using Stirling's approximation for $k!$ and writing

$$(22) \quad |\phi(x)| \leq \sigma^k \delta^k k^{rk} \|f\|_h$$

where σ is a constant multiple of ρ and, what is important, independent of all other parameters under consideration.

Now comes the tricky part. Let

$$\omega = \max \left\{ \sigma, \left(\frac{2}{3} \right)^r \frac{2}{b_0} \right\} \text{ and } \delta_0 = \frac{1}{3^r \omega (m+1)^r}.$$

Choose k so that $1 \leq 3(\omega\delta)^r k \leq 2$. These choices imply that $k \geq m+1$ and, since $r \geq 1$, $2\delta k \leq 2\delta k^r \leq b_0$, so that all the calculations leading to (22) are valid. Furthermore, $(\sigma\delta)^{1/r} k \leq 2/3$ and $k \geq 1/(3(\omega\delta)^{1/r})$ so that (22) implies (15) with $\lambda = (2/3)^{1/(3\omega^{1/r})}$. This completes the proof of (15).

To see (16) assume $0 \leq r < 1$, let σ be the constant in (22), and set

$$\delta_1 = \min \left\{ \frac{b_0}{4(m+1)}, \left(\frac{1}{\sigma} \left(\frac{2}{b_0} \right)^r \right)^{1/(1-r)} \right\}.$$

Suppose δ is a positive number $\leq \delta_1$ and take the integer k so that it satisfies $b_0/2 \leq 2\delta k \leq b_0$. This choice of k ensures that $k \geq m+1$ and $2\delta k \leq b_0$ so that all the calculations leading to (22) are valid. Furthermore, $k \leq b_0/(2\delta)$ so that

$$(\sigma^k k^r)^k \leq \left\{ \sigma \delta \left(\frac{b_0}{2\delta} \right)^r \right\}^k$$

and since the expression in braces is ≤ 1 and $k \geq b_0/(4\delta)$ we may write

$$\left\{ \sigma \delta \left(\frac{b_0}{2\delta} \right)^r \right\}^k \leq \left\{ \sigma \delta \left(\frac{b_0}{2\delta} \right)^r \right\}^{b_0/(4\delta)} = \left\{ \left(\frac{\sigma b_0^r}{2^r} \right)^{1/(1-r)} \delta \right\}^{(1-r)b_0/(4\delta)}.$$

Finally, in view of (22) and the above inequalities, we may conclude that if $0 \leq \tau < 1$ and $0 < \delta \leq \delta_1$ then (16) holds with

$$C = \left(\frac{\sigma b_0^r}{2^r} \right)^{1/(1-\tau)} \quad \text{and} \quad c = \frac{(1-\tau)b_0}{4}.$$

■

We draw the reader's attention to the following points concerning these theorems:

- The hypotheses imply that X is unisolvent for \mathcal{P}_k .
- In view of well known estimate on $E_k(h, \epsilon)$, Theorem 1 implies that error bound is $O(\delta^{(k+1)/2})$ as $\delta \rightarrow 0$ for arbitrarily large k whenever h is sufficiently smooth.
- The cube Ω can be replaced by more general regions. For example, by using affine transformations it should be clear that Ω can be taken to be a parallelepiped.
- Consider (14) and (16) in Theorem 2. In the case $\tau = 0$ the measure μ has compact support and so do the members of \mathcal{C}_k . In this case the constant C in (16) can be taken to be independent of b_0 and c can be taken to be $b_0/4$. Taking $b_0 = \infty$, for example this is the case whenever Ω contains an open cone, it follows that $s(x) = f(x)$. This means, of course, that the values of f on X uniquely determine f .
- The estimate needed for n -variate polynomials to replace the estimate on $\sum |\ell_j(x)|$ in the above argument may be found in [8]. The remaining ingredients needed for the proof in the general n -variate case can be found in [7].

4 Miscellaneous examples and remarks

Observe that the examples mentioned in the introduction $h(x) = e^{-|x|^2}$, $(1 + |x|^2)^{-(n+1)/2}$, and $-\sqrt{1 + |x|^2}$ are very smooth so that $E_k(h, \epsilon) = O(\epsilon^{k+1})$ for any non-negative integer k in all these cases. The first example satisfies (14) with $\tau = 1/2$ and the other two satisfy (14) with $\tau = 1$. For more examples of this type see [7].

The function h defined by

$$h(x) = \prod_{j=1}^n \frac{\sin x_j}{x_j}$$

where x_1, \dots, x_n are the coordinates of x , in other words $x = (x_1, \dots, x_n)$, is conditionally positive definite of order zero and satisfies (14) with $\tau = 0$ and $\rho = \sqrt{n}$.

The other example mentioned in the introduction $h(x) = -|x|$ is not so smooth and fails to satisfy (14). More generally we may consider $h(x) = |x|^a$, $a > 0$, which fail to satisfy (14) but enjoy $E_k(h, \epsilon) = O(\epsilon^a)$ as $\epsilon \rightarrow 0$ whenever $k \geq a - 1$. Thus Theorem 1 implies error bounds which are $O(\delta^{a/2})$ as $\delta \rightarrow 0$.

for these examples. Such error bounds are not optimal. In the case $n = 1$ and $a = 1$ or 3 the corresponding interpolants are piecewise linear or piecewise cubic splines respectively; the optimal pointwise error estimates are known to be $O(\delta^{a+1})$ as $\delta \rightarrow 0$ in these instances, see [11]. In the case $n = 3$ and $a = 1$ Nelson [4] has shown that a pointwise error bound is possible which is $O(\delta^4)$ as $\delta \rightarrow 0$.

The variational theory of h -splines introduced in [5] may be regarded as an extension of [2]. The motivation for this extension was a question implicitly raised in [3] concerning the invertibility of $(a_{ij}) = (\sqrt{1 + |x_i - x_j|})$ which this theory settled. An alternate, apparently more appealing, solution to this question is Micchelli's generalization [9] of Schoenberg's theorem [10] concerning positive definite functions; this extension inspired many related publications which, because of space limitations, cannot be cited here, however, see the related articles by Buhmann, Powell, Schaback, and Ward in these proceedings.

References

- [1] P. J. Davis, *Interpolation and Approximation*, Dover, New York, 1975.
- [2] J. Duchon, Splines Minimizing Rotation-invariant Semi-norms in Sobolev Spaces, *Constructive Theory of Functions of Several Variables*, Oberwolfach 1976, W. Schempp and K. Zeller, ed., Springer-Verlag, Berlin-Heidelberg, 1977, p. 85-100.
- [3] R. Franke, Scattered data interpolation: tests of some methods, *Mathematics of Computation* 38 (1982), 181-199.
- [4] R. L. Hardy and S. A. Nelson, Approximation of Newtonian Potentials, *Approximation Theory V*, Chui, Schumaker, Ward, eds., Academic Press, Orlando, 1986, 375-378.
- [5] W. R. Madych and S. A. Nelson, Multivariate interpolation: a variational theory, Technical Report, 1983.
- [6] W. R. Madych and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions, *Approx. Theory and its Applications* 4 (1988) 77-89.
- [7] W. R. Madych and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions II, *Math. Comp.* 54 (1990) 211-230.
- [8] W. R. Madych and S. A. Nelson, Bounds on multivariate polynomials and exponential error estimates for interpolation via generalized splines, BRC Technical Report, 1990.
- [9] C. A. Micchelli, Interpolation of scattered data: distance matrices and conditionally positive definite functions, *Constr. Approx.* 2 (1986) 11-22.
- [10] I. J. Schoenberg, Metric spaces and completely monotone functions, *Ann. of Math.* 39 (1938), 811-841.
- [11] L. L. Schumaker, *Spline Functions*, Wiley, New York, 1981.